# The Firm as an Entity

## Implications for economics, accounting and the law

*Edited by*

**Yuri Biondi, Arnaldo Canziani and Thierry Kirat**

The Economics of Legal Relationships

# The Firm as an Entity

The recent crises in corporate governance and disclosure have highlighted the deficiencies and shortcomings of received approaches to the theory of the firm – the managerial and enforcement recommendations they advocate imply poor policy and sometimes huge economic and societal losses. This suggests the need for a new synthesis.

Through new essays and valuable reprints, a host of eminent international contributors seek to provide new insights into the theory of the firm, utilising an interdisciplinary approach that integrates law and economics with accounting. This new perspective links the three disciplines through the synthetic notion of the firm as an entity, and wishes to create a deeper understanding of the organisational and institutional dimensions of the firm.

*The Firm as an Entity* will provide the reader with (i) general essays, that pay attention to these insightful lessons from the past that presently are at least partly neglected; (ii) specific essays, that contribute to the enhancement of the interdisciplinary perspective by exploring its consequences and implications for accounting, ontology, law and economics, business finance, and the governance of the firm as an enterprise entity; and (iii) reprints of often rare or neglected essays, by H. Simon, M. Shubik, R. Coase, A.A. Berle Jr., R.N. Anthony and J.H. Stauss, that allow the reader to rediscover the work of previous authors, issues and ideas, and that will prove to be fundamental to enhance the interdisciplinary approach offered by the collection.

**Yuri Biondi** is Associate Professor, Conservatoire National des Arts et Métiers (CNAM), Paris, France.

**Arnaldo Canziani** is Professor of Economia Aziendale, Faculty of Law, University of Brescia, Italy.

**Thierry Kirat** is Research Fellow, Centre National de la Recherche Scientifique (CNRS), France.

## The Economics of Legal Relationships

Series editors:
Nicholas Mercuro and Michael D. Kaplowitz, Michigan State University

**Compensation for Regulatory Takings**
*Thomas J. Miceli and Kathleen Segerson*

**Dispute Resolution**
Bridging the settlement gap
*Edited by David A. Anderson*

**The Law and Economics of Development**
*Edited by Edgardo Buscaglia, William Ratliff and Robert Cooter*

**Fundamental Interrelationships Between Government and Property**
*Edited by Nicholas Mercuro and Warren J. Samuels*

**Property Rights, Economics, and the Environment**
*Edited by Michael Kaplowitz*

**Law and Economics in Civil Law Countries**
*Edited by Thierry Kirat and Bruno Deffains*

**The End of Natural Monopoly**
Deregulation and competition in the electric power industry
*Edited by Peter Z. Grossman and Daniel H. Cole*

**Just Exchange**
A theory of contract
*F.H. Buckley*

**Network Access, Regulation and Antitrust**
*Edited by Diana L. Moss*

**Property Rights Dynamics**
A law and economics perspective
*Edited by Donatella Porrini and Giovanni Ramello*

**The Firm as an Entity**
Implications for economics, accounting and the law
*Edited by Yuri Biondi, Arnaldo Canziani and Thierry Kirat*

**The Legal-Economic Nexus**
*Warren J. Samuels*

The first three volumes listed above are published by and available from Elsevier.

**The Economics of Legal Relationships**

*Sponsored by the Michigan State University College of Law*

*STATEMENT OF SCOPE*

*The Economics of Legal Relationships* is a book series dedicated to publishing original scholarly contributions that systematically analyze legal-economic issues. Each book can take a variety of forms:

- It may be comprised of a collection of original articles devoted to a single theme, edited by a guest volume editor.
- It may be a collection of refereed articles derived from the Series Editors' "call for papers' on a particular legal-economic topic.
- An individual may wish to (co)author an entire volume.

Each book in the series is published in hardback, approximately 250–300 pages in length and is dedicated to:

- Formulate and/or critique alternative theories of law and economics – including – the new law and economics, the economics of property rights, institutionalist and neoinstitutionalist law and economics, public choice theory, Austrian law and economics, or social norms in law and economics.
- Analyze a variety of public policy issues related to the interface between judicial decisions and/or statutory law and the economy.
- Explore the economic impact of political and legal changes brought on by new technologies and/or environmental and natural resource concerns.
- Examine the broad array of legal/economic issues surrounding the deregulation/re-regulation phenomena.
- Analyze the systematic effects of legal change on incentives and economic performance.

*CALL FOR AUTHORS / VOLUME EDITORS / TOPICS*

An individual who is interested in either authoring an entire volume, or editing a future volume of *The Economics of Legal Relationships* should submit a 3–5 page prospectus to one of the series editors. Each prospectus must include: (1) the prospective title of the volume; (2) a brief description of the organizing theme of the volume whether single authored or edited; (3) an identification of the line of literature from which the proposed topic emanates, and (4) either a table of contents or, if edited, a list of potential contributors along with tentative titles of their contributions. Send prospectus to either series editor. Please note that the series editors only accepts individual manuscripts for publication consideration in response to a specific "Call for Papers."

**Send prospectus directly to either series editor:**

Professor Nicholas Mercuro
Michigan State University
College of Law
East Lansing, MI 48824
phone: (517) 432-6897
e-mail mercuro@msu.edu

Professor Michael D. Kaplowitz
Michigan State University
Department of Community Agriculture, Recreation, and Resource Studies
East Lansing, MI 48824
phone: (517) 355-0101
e-mail kaplowit@msu.edu

# The Firm as an Entity

## Implications for economics, accounting and the law

**Edited by**
**Yuri Biondi, Arnaldo Canziani and Thierry Kirat**

Routledge
Taylor & Francis Group
LONDON AND NEW YORK

First published 2007
by Routledge
2 Park Square, Milton Park, Abingdon, Oxon OX14 4RN

Simultaneously published in the USA and Canada
by Routledge
270 Madison Ave, New York, NY 10016

*Routledge is an imprint of the Taylor & Francis Group, an informa business*

This edition published in the Taylor & Francis e-Library, 2008.

"To purchase your own copy of this or any of Taylor & Francis or Routledge's collection of thousands of eBooks please go to www.eBookstore.tandf.co.uk."

© 2007 Editorial and selection matter, Yuri Biondi, Arnaldo Canziani and Thierry Kirat; individual chapters, the contributors

All rights reserved. No part of this book may be reprinted or reproduced or utilized in any form or by any electronic, mechanical, or other means, now known or hereafter invented, including photocopying and recording, or in any information storage or retrieval system, without permission in writing from the publishers.

*British Library Cataloguing in Publication Data*
A catalogue record for this book is available from the British Library

*Library of Congress Cataloging-in-Publication Data*
A catalog record for this book has been requested

ISBN 0-203-93111-4 Master e-book ISBN

ISBN 978–0–415–41443–2 (Print Edition)

# Contents

*List of illustrations* x
*List of contributors* xi

**PART I**
**Introduction: the firm as an entity** **1**

1 Coming back to the enterprise entity 3
YURI BIONDI, ARNALDO CANZIANI AND THIERRY KIRAT

2 The economic theory of the firm as an entity: an overview of the volume 9
YURI BIONDI

**PART II**
**On the economic theory of the firm as an institution and an organisation** **19**

3 The current state of the economic theory of the firm: contractual, competence-based, and beyond 21
OLIVIER WEINSTEIN

4 Organisations and markets 54
HERBERT A. SIMON (1991)

5 Accounting and its relationship to general equilibrium theory 73
MARTIN SHUBIK (1993)

6 Accounting and the theory of the firm 82
RONALD H. COASE (1990)

7 The impact of the corporation on classical economic theory 92
ADOLF A. BERLE JR (1965)

**PART III**
**Perspectives for accounting, law and economics: lessons from the past** **105**

8 *Economia aziendale* and *Betriebswirtschaftslehre* as autonomous sciences of the firm 107
ARNALDO CANZIANI

9 The firm between law and economics: an overview of selected legal-economic scholars of the past 131
THIERRY KIRAT

10 A historical perspective on corporate form and real entity: implications for corporate social responsibility 153
REUVEN S. AVI-YONAH AND DGANIT SIVAN

11 The theory of enterprise entity 186
ADOLF A. BERLE JR (1947)

12 The trouble with profit maximization 201
ROBERT N. ANTHONY (1960)

13 The entrepreneur: the firm 216
JAMES H. STAUSS (1944)

**PART IV**
**Essays on economic, legal and accounting features of the firm as an entity** **235**

14 Accounting and the economic nature of the firm as an entity 237
YURI BIONDI

15 Some building blocks for a theory of the firm as a real entity 266
DAVID GINDIS

16 The firm as an entity and the law: the economic substance, the legal forms 292
FEDERICO MANFRIN

17 Economics and finance of the firm as an entity 317
GIUSEPPE MARZO

18 The corporate governance of the firm as an entity: old issues for the new debate 348
MARC T. MOORE AND ANTOINE REBÉRIOUX

*Index* 375

# Illustrations

## Figures


12.1 Example of necessary estimates of demand and costs 206
14.1 The usual dualism between the firm and the market 240
14.2 The firm-entity as an institution and an economic organisation 254
17.1 Financial needs and provisions 341
17.2 Temporary and permanent provisions 342


## Tables


10.1 Theories of the corporation and corporate social responsibility (CSR) 174
14.1 Equilibrium and accounting frameworks for the economic and monetary process of the firm 249
14.2 Proprietary and entity views of the firm 251

# Contributors

**Reuven S. Avi-Yonah** is the Irwin I. Cohn Professor of Law at the University of Michigan.

**Yuri Biondi** has worked on accounting and the economics of non-business entities in education, on the foundations of the Italian tradition of business economics and accounting, and on the economic nature of the business firm as an entity between economic theories and accounting. He is currently associate professor at the Conservatoire National des Arts et Métiers (CNAM, Paris).

**Arnaldo Canziani**, professor of *Economia Aziendale* (business economics) at the University of Brescia. His research interests include the Italian and Continental tradition of accounting and business economics.

**David Gindis** is a lecturer in economics at the Institut National des Sciences Appliquées (INSA) in Lyon, and is currently completing a PhD dissertation on the economics of the firm at the University of Lyon II. His research interests include the theory of the firm, law and economics, institutional economics and philosophy of economics.

**Thierry Kirat** is Centre National de la Recherche Scientifique (CNRS) Research Fellow on Law and Economics issues, attached to the Institut de Recherche Interdisciplinaire en Sociologie, Économie, Science politique (IRISES), University Paris-Dauphine.

**Federico Manfrin** has been lecturer of Criminal Law at Brescia University (2003–5) and Administrative Law at Milan University (2004–5). He is now assistant Researcher in Economics and Law at Brescia University.

**Giuseppe Marzo** is researcher in Economia Aziendale (Business Economics) at the University of Brescia, and a lecturer in Business Strategy and Firm Valuation at the University of Ferrara, and in Business Communication at the Università Cattolica del Sacro Cuore in Brescia. His research interests are in the fields of intangible assets and real options analysis and valuation.

**Marc Moore** is a lecturer in law at the University of Bristol, UK. His research interests are in corporate governance and financial markets.

**Antoine Rebérioux** is associate professor of economics at the University of Paris X (Nanterre) and a member of *EconomiX*, a research unit attached to the CNRS. He works on the theory of the firm and corporate governance, and has recently coauthored (with M. Aglietta) the essay "Corporate Governance Adrift. A Critique of Shareholder Value" (Edward Elgar Publishing, 2005).

**Dganit Sivan** is an SJD candidate at the University of Michigan. She obtained her LLM there in 2005.

**Olivier Weinstein** is Professor in Economics at the University Paris 13 and a member of the CEPN (Centre d'Economie de Paris Nord), a research unit attached to the CNRS. His areas of interest are the theory of the firm, the economy of institutions and of organisations, the institutional approaches of markets, and the economics of innovation and R&D.

# Part I

# Introduction

## The firm as an entity

# 1 Coming back to the enterprise entity

*By Y. Biondi, A. Canziani, and T. Kirat (volume editors)*

## A feeling of lack about the current state of the economic theory of the firm

Notwithstanding their insights, and their innovative contributions to a better understanding of the economy of the firm, the accepted theories of the firm still have a flavour that is distinctly neoclassical, i.e., they directly or indirectly rely on "worldviews" (*Weltanschauungen*) as well as on tools rooted in the neo-classical tradition. Such a common approach rests on two basic preconceptions. Drawing upon methodological individualism, which reduces collective phenomena to transient interactions among individuals, the collective dimension of the firm is entirely associated to the question of the legal personality and thus to the corporation. As a consequence, once the firm – seen as a corporation – dissolves in atomistic interplays, or is merely understood either as a "legal fiction" or a "governance device" dealing with the relations among stakeholders, it loses its fundamental contents of an economic, holistic and dynamic nature.

As a matter of fact, the theoretical and analytical interest in the firm is not at all new, but it is far more long-standing than generally imagined. It clearly emerged at the beginning of the twentieth century when scholars – in the United States, and mainly in Continental Europe (France, Germany, Italy) – realised that economic transformations under way asked for innovative inquiring about the firm and its active impact on economy and society. In addition, the need for new perspectives was fostered by the desire to delve into facts, and by the deeper insight driven by epistemic revolutions as well, pragmatism to mention but one.

Economic theory actually retained some of the contributions of that period – those concerned in particular with oligopolies and limited competition in general – but the entire movement was broad and insightful. The economic institutions of the renewed industrial and financial capitalism (big business, i.e. groups and trusts and *Konzerne*) radically challenged the classical theories as well as the interpretation of economic and social relationships they spread. Generally speaking, the whole of received perspectives – with their deterministic, mechanical, reassuring appeal – were under attack: questioned and defied. From this scientific unrest, a set of theoretical and practical issues arose: (i) dynamics

and the firm (or enterprise); (ii) the nature and role of profit; (iii) the nature and role of the firm and its key actor (be she proprietor, entrepreneur, manager, captain of industry, "baron"); (iv) money, production and value.

## The comprehensive approaches to the world of firms in some neglected authors of the past

Referring to that time of high theory, we revive here two different sources of approaches and ideas: (i) American institutionalism; and (ii) the continental tradition of accounting and business economics, coupled with the seminal contribution of A.C. Littleton in the United States. At the same time, we cannot neglect a galaxy of scholars from different nations and fields – from W. Rathenau to F. Perroux – who were sharing common scientific concerns and adopting broadly speaking institutional perspectives.

Challenged by the prophetical essays of T. Veblen on business enterprise, absentee ownership and the leisure class, American institutionalism focused the economic and monetary process generated by business dynamics. John R. Commons in parallel, along with the transformation of American economy, masterfully interpreted the interplay of two core institutions of capitalism, law and money, especially thanks to impressive empirical analyses and original theorisation. He centred on the institutional economic features of this process and on the ways that should lead to a reasonable functioning of the capitalistic system. At the same time, John Maurice Clark devoted much of his lifetime to study the social control of business. His analyses of the transformation of American capitalism were dominated by the problem of accounting for overheads, and the main features of institutional processes.

With the valuable collaboration of Means, A.A. Berle Jr developed a paramount research programme on economics, finance and law that generated the well-known, seminal analysis of the role of modern corporations and private property. According to Berle, the economic theory of the twentieth century needed definite severance from classical economics, in particular from the idea of property as ownership of the business enterprise as well as the very defender of liberty, justice and peace. The classical framework – especially private property rights – is no longer able to cope with the whole congeries of the firm, whose dynamic and collective dimension open the way and constitute the field of overwhelming power ; a power that, in turn, theory and polity have to grasp and deal with, in order to make liberty and fairness effective.

Veblen, Commons and Berle stand as the leading figures among a generation of economists who cast doubts on both Classical (English) and emerging neoclassical perspectives, providing the more radical tenets of marginalism have been influential for Commons or even Zappa, in an early phase. Their contributions were fostered by a common quest, which, in turn, needed to be understood. Gruchy (1947) – describing their economic thought by the term "holistic" – worked out the first synthesis of this relevant, coherent alternative to classical and neoclassical perspectives.

This institutional economics "calls for the construction of an economic science that would pay more attention to the theory of production and less to the mechanics of a competitive price system" (Gruchy 1947: 7). As a consequence, the old economics of the market recedes into the background, while attention turns to such problems as surplus or inequalities of income distribution, i.e., to some of the relevant deficiencies of the competitive price system. Such an institutional economics adopts a comprehensive worldview in looking to reality and interpreting it. Economy and society are seen as an integrated system, as an evolving pattern or complex with a past, present and future. This system cannot be understood by splitting it into parts: collective and dynamic characters enter the drama. This is why the competitive price system, centred on equilibrium, is no longer helpful, nor as first approximation: whenever an economic approach relies on methodological reductionism, it appears to be essentially atomistic and static. On the contrary, the system presents an essential wholeness: not a mere assemblage of detached parts and elements but – as a matter of fact – an interrelated, structured whole whose parts are dependent and complement each other. This does not mean that all parts of the whole have a necessary harmony between them, but it implies a kind of coordination that is totally different from the equilibrium pattern provided by a formal, stable, determined mechanism. The whole is brought about by a number of factors which imply principles and choices, purposes and constraints, order and disorder, efficiency and waste, fairness and unfairness, development or distress: open to the forces of economic and social change, such a whole results in being only relatively established and endlessly evolving.

American scholars were not alone in this quest. At the very same time in Continental Europe, namely in Germany and Italy, business economists and accountants such as Nicklisch, Schmalenbach and Zappa developed self-contained theories of the firm. These theories moved from the firm interpreted as an economic agency to produce wealth, progressed steadily along time considering the firm's social roles and interplays (fulfilling human needs, producing and distributing incomes, playing a distinctive social role), and identified it at the end as an institutional construct. This renewed interpretation contrasted among others the "black box" view: mechanical and old-fashioned, but still influential.

The two aforementioned streams, as well as the galaxy of authors adopting an institutional tenet, crossed the subsequent decades in either a well-recognised or a hidden way. And just to refer here to the most important of them from the viewpoint of continuity and coherence (Herbert Simon apart), François Perroux (1966) launched a comprehensive inquiry on "The Firm and the Economics of Twentieth Century." He asked Berle and Richardson, among others,[1] to write about the firm, its functions and role in the new context of the late twentieth century. In his introduction to the collection, Perroux spoke of the limits of a "purely market theory" – i.e., the neoclassical black box – in understanding the economy of the firm. In particular, moving from market failures, and contrary to the mainstream, he posed such key questions as what are the intrinsic

deficiencies of the price system as well as its lack of understanding economic relations and interactions under real dynamics and complexity.

At the same time, moving from the idea of "bounded rationality," Simon (1991; 1997) developed a valuable critique of the firm framed by equilibrium, and suggested interpreting the firm as a dynamic system, connecting himself to Commons and the American institutionalism as well. Therefore, Simon stressed the need for a new theory grounded on *active* firms instead of paramount "efficient markets," and criticised both the profit maximisation approach and the underlying idea of neoclassical equilibrium – essentially a static one.

## The concept of the firm as an entity

Evidently both Simon and Perroux share the opinion that delving into the reality of the whole firm helps to recognise the limits of a purely market theory of the firm. Only in its abstract world can the firm as a "black box" provide the rationale for understanding and managing the firm. Under real dynamics and complexity, on the contrary, no simplistic solution can substitute for the interplay of management, organisation and accounting systems. No *mechanism* deprived of heuristic content can substitute for the idea and role of responsible decision-making, since the firm is a *dynamic system of interactions, interdependencies, and complementarities*. These relationships, located in time and space, are not merely (nor fully) contractual nor bargained: their dynamic system is different in nature from any static equilibrium of prices or nexus of contracts, property rights included.

These features are some of the building blocks of a theory of the firm as an entity. Anyway, the same concept of entity appeared first within the accounting field, to identify the core of the business activity which costs and revenues, financing and investments can be attributed to. Constructing on the many similar insights of those authors, and on Zappa's in particular, the economic theory of the firm as an entity may be understood as the *actual economic coordination* set up by the management system – especially through the implementation of a working organisation – coping with the *ongoing economic process* that accounting represents and helps to govern.

This way, the concept of entity appears to constitute the innovative, unifying core of those theories of the firm. Such a new framework allowed the definition of new research territories as well as new approaches to economic and financial dynamics, income to the firm, and the production and distribution of wealth.

The very innovation of this perspective consists anyway in the different way through which these territories are entered into: contrary to methodological individualism, a comprehensive framework is developed, consisting in realism, wholeness and dynamics. At the very same time, the concept of entity permits the synthesis between the concepts of *whole* proper to institutional economics and of *dynamic system* proper to Continental accounting and business economics; and integrates the two concepts of "*Betrieb*" and "*azienda*," originally accounting ones.

Due to these (and other) reasons, this theory of the firm appears to deserve consideration as it permits the everlasting juxtaposition between theories without facts and facts without theories to be overcome.

In particular, it allows the economic analysis of the firm to be enriched by accounting and law. Accounting provides a better understanding of the economic and financial dynamics of the firm, driving to the very comprehension of such matters as assets, liabilities, revenues, prices, transactions, as well as to the appropriate representation of them. Legal discourse developed a view of the firm as an entity distinct from its mere legal form: some features differed due to legal tradition; anyway they were not so far from each other as regards some basic underpinnings. In addition, the world of law shapes the reality of economic and financial dynamics in an essential way, contrary to the ajuridical approach of most theories of the firm.

Further to our sharing of the aforementioned frameworks, the following collective effort offered to the readers aims at enriching the economic theory of the firm with such an interdisciplinary approach that integrates law and economics with accounting, while the three fields are linked by the synthetic notion of the firm as an entity.

As a conclusion, we hope this book can set forth a contribution to the theory of the firm, based on this interdisciplinary notion and focusing on its accounting, legal and economic features. The proposed view of the firm – entity, whole, and dynamic system as well – comes both from some new essays, and from rediscovering the works, concerns and ideas of authors of the past whose reprints appeared essential to ground the approach offered here.

## Notes

1 Here is the complete list of contributors in order of appearance in the three volumes. First volume: F. Perroux, F. Bloch-Lainé, M.A. Adelman, W.J. Baumol, M. Demonque, D. Chandler Jr and S. Salsbury, C. O'Donnell, N.W. Chamberlain, J. Pajestka and K. Porwit, S. Lombardini, J. Messner, H. Chambre, J. Houssiaux. Second volume: H. Theil, M.K. Starr, L. Devaux, G.A. Steiner, R. Faure and A. Kaufmann, P. Laffitte, Ch. Ribet-Petersen, E.P. Learned, A. Touraine, R.J. Monsen, B.O. Saxberg, R.A. Sutermeister, G. Lasserre, M. Maire, F.R. Wickert, J. Lesourne. Third volume: A.A. Berle, J. Dean and W. Smith, R. Harris, P. de Woot, A. Nowicki, C.W. Churman, R.M. Cybert and L.B. Lave, G.B. Richardson, B.S. Keirstead, H. Koontz, S.P. Dobrovolski, F. Perroux.
This inquiry anticipated by few years the research impulse given by Coase (1972) through his NBER conference on industrial organisation (University of Chicago, November 1970), which certainly contributed to the rediscovery of his article on the nature of the firm that Perroux (1966: 15, note 1) had already mentioned and acknowledged.

## References

Coase, Ronald (1972), "Industrial Organization: A Proposal for Research," in Victor R. Fuchs, ed., *Policy Issues and Research Opportunities in Industrial Organization*. New York:

NBER, 59–73. Reprinted in *The Firm, the Market and the Law*, University of Chicago Press, 1988, 57–74.

Commons, R. John (1924), *Legal Foundations of Capitalism*, USA: Macmillan. Reprinted USA: Transactions Pub. 1996.

—— (1934), *Institutional Economics. Its Place in Political Economy*, USA: Macmillan. Reprinted USA: Transactions Pub. 1990.

Gruchy, G. Allen (1947), *Modern Economic Thought*, New York: Prentice-Hall Inc.

Littleton, A. Charles (1953), *Structure of Accounting Theory*, USA: AAA.

Nicklisch, Heinrich (1932), *Die Betriebwirtschaft* [The Economics of Enterprise Entity], 7th edn, Stuttgart: C.E. Poeschel Verlag. Reprinted Glashuetten im Taunus: Verlag Detlev Auvermann KG, 1972.

Perroux, François (1966), “L’entreprise et le phénomène collectif du second XXe siècle” [The firm and the collective phenomenon of the late twentieth century], vol. I, pp 1–17, in F. Bloch-Lainé and F. Perroux (eds), *L’entreprise et l’économie du XXe siècle. Étude internationale* [The Firm and the Economics of Twentieth Century. An International Inquiry], three vols, Paris: PUF.

Rathenau, Walter (1918), *Von kommenden Dingen*, Berlin, trans. E. and C. Paul, “In Days to Come,” London 1921.

Schmalenbach, Eugen (1926), *Dynamische Bilanz*, 3rd edn, Leipzig. English translation: *Dynamic Accounting*, London: Gee and Co. 1959.

Simon, A. Herbert (1991), “Organizations and markets,” *Journal of Economic Perspectives*, 5 (2), spring: 25–44, reprinted in this volume.

—— (1997), *An Empirically Based Microeconomics*, Cambridge: Cambridge University Press.

Veblen, Thornstein (1899), *The Theory of the Leisure Class: An Economic Study in the Evolution of Institutions*, New York and London: Macmillan, Feb. 1899. Subtitle changed in 1912 to: *An Economic Study of Institutions*.

—— (1904), *The Theory of Business Enterprise*, New York: Scribner’s. Reprinted: Transaction Books, New Jersey (USA): 1978.

—— (1923), *Absentee Ownership*, B.W. Huebsch, New York: 1954.

Zappa, Gino (1937), *Il reddito d’impresa* [The Income to the Firm], Milan: Giuffrè.

# 2 The economic theory of the firm as an entity

## An overview of the volume

*Yuri Biondi*

### The structure of the volume

The collective book we offer to the reader seeks to provide a new contribution to the theory of the firm based on the interdisciplinary synthetic notion of entity through new essays and reprints of the work of relevant past scholars. This overview summarises the volume and provides the reader with a first instructive guideline. It thus sets aside the peculiarities of each paper in order to exhibit their complementarity in jointly achieving the broader scope of the collection as a whole.

The volume comprises an introductory part followed by three different parts.

The first and introductory part (the present chapter included) summarises the joint intent of the editors and authors by presenting the idea of entity and the scope of the book.

The second part deals with current approaches to the firm (Weinstein) and summarises some key issues fundamentally open in the current state of the theory of the firm. These points are especially stressed by reprints from leading scholars such as Simon, Shubik and Coase, which together call for a new perspective capable of better understanding the institutional and organisational dimensions of firms.

In fact, some great scholars have already raised these issues and suggested challenging ways to understand the business firm and its impact on economy and society. Their approaches and perspectives merit continued consideration. Indeed, the third part deals with that time of high theory that presently is at least partially neglected. The original essays synthesise the major themes with which the volume is concerned (Canziani; Kirat; Avi-Yonah and Sivan), while other reprints offer valuable insights driven from the past. Present and past contributions together offer some building blocks for a view of the firm as an entity, a whole, a dynamic system. In particular, such a view integrates law and economics with accounting. The firm is then seen as an enterprise entity, characterised by a peculiar economic and monetary process that generates a special income while being confronted with the actual dynamics and complexity of reality.

Drawing and expanding upon the first three parts, the fourth part comprises new essays on special features of the firm seen as an entity. This notion constitutes a new way to enter the organisational and institutional dimensions of the firm as

a whole and a dynamic system, and each chapter contributes to the enhancement of the interdisciplinary perspective by exploring its consequences and implications for accounting (Biondi), ontology (Gindis), law and economics (Manfrin), business finance (Marzo), and the governance of the enterprise entity (Moore and Rebérioux).

## Overview of the second part

Weinstein opens the second part with a broad and meaningful reconstruction of the new theories of the firm. His purpose is not to present an exhaustive overview, but to point to certain lessons and questions concerning the current understanding of the firm, its realities and the possible principles of its economic theorisation. Within the contractarian perspective, Weinstein distinguishes two different patterns, starting from Coase's or Berle and Means' legacies. Moreover, distinguishing contractarian and competence-based perspectives, and starting from the advances these approaches provide, Weinstein disentangles the issues that open to a possible synthesis. He underlines the limits of a view based only on contracts or individualistic property rights, and stresses the relevant insights provided by comprehensive approaches that consider the special process of production generated by the whole firm. The latter, building on Chandler and Simon, appear to be the most promising for a new perspective capable of understanding the firm as a real entity.

As Weinstein explains, the more recent debate stresses the "incompleteness of contracts" as the key feature that shapes the economic organisation of firms. Inside the firm, such incompleteness reveals the matters that make contractual analysis unable to comprehend the firm's whole economic organisation; matters of relation, dynamics, implied and shifting conditions, and context.

The abstract world of sole contracts (property rights included) can be reduced to a universe of sole individuals who, following their own interest, are governed in their choices by a system of prices in equilibrium. Their contractual relationships are always featured by the capacity to not enter or to exit the relation, and such a purely market process is then characterised by the independent coexistence of free and equal individuals governed by the "invisible hand."

Its political and ethical implications apart, once theory engages in confronting reality it enters an entirely different world, shaped by real dynamics and complexity, and characterized by discretion and power driven by asymmetries, inequalities and changes. Contracts represent at best a part of the real drama, sometimes a merely formal one. In this positive context, understanding contractual incompleteness requires the reconstruction of the notion of contract itself. In the real world fraught with incompleteness, as one past scholar said, "everything in the contract is not contractual" (Durkheim 1930: 189). As it did at that time of high theory, theory is still confronted with the "socialisation" of firms (and other relevant socio-economic activities) that makes a purely market view out of date. Instead of the "invisible hand," theory has to recognise at least the visible hand of management and actual institutional arrangements. The exit

right then becomes just a possibility offered by contract law, but the contract itself relates to a whole of claims and obligations imperfectly enforced by institutions that prompt, frame and enhance them.

Once the incompleteness of contracts becomes one of the key features in the real world, this further implies the abandonment of a world of sole contracts to enter a world of organisations, institutions and dynamics. In the latter world, contract may be reconsidered starting from its inter-personal nature. It then involves a socio-economic relation, and may give a form to this relation that *makes society* between actors and individuals concerned with real purposes and needs. A comprehensive view of the economic and monetary process may start from this whole of meaning, relations and interactions, contracts included, that allows actors and individuals to jointly purport and intend. A change of perspective is required. The kind of change Simon (1991) suggests by his Martian view.

By dealing with organisations and markets, Simon (1991) raises the challenge of understanding such an "organisational economy." He criticises the shadowy firm entirely framed by a system of prices in equilibrium, claiming against praised lack of realism in assumptions (Simon 1979). Starting from both a realistic tenet and a theoretical connection with Commons and old American institutionalism, Simon calls for a new theory grounded on *active* firms instead of on paramount efficient markets. Such a theory may deal with the firm as a dynamic system, situated in time and space, and guided by a *satisficing* principle of continuity handling with bounded rationality. Therefore, the equilibrium machinery that justifies the profit maximisation principle has to be replaced by increased focus on the actual knowledge-generating systems that management uses to cope with the special economic environment generated by the firm. In the third part, Anthony (1960) further develops this point by a dynamic accounting contribution.

To enter the real world of firms and organisations, a further step is then required, i.e., the critique of the usual equilibrium approach. In looking for the real factors affecting the process of production and exchange, a sharp line exists between formalisation as an instrument of interpretation of reality, or as the achievement of an imaginary world abstracted away from any heuristic content. Shubik (1993) stigmatises the formal limits of equilibrium economics by exploring the pragmatic contribution of accounting to economic theory. Accounting is here both a fact and a practical viewpoint that theory has to properly deal with.

Equilibrium economics reduces the firm to a timeless shallow nexus of disparate parts or individuals alone. It neglects the genuine implications both of the actual complexity of the whole firm, and of its economic and monetary process confronted with real time and uncertainties. According to Shubik (1993), time and uncertainties have essentially disappeared from the apotheosis of the *system of pricesin equilibrium*, but they remain the concerns of everyday business activity. The problems of how to account for their influence in the *ongoing economic process* are central to the development of accounting. They lead accounting to understand the firm and its special economy as an entity, i.e., as a dynamic system that the accounting system helps to represent and govern. The accounting logic

and structure based on the entity view are designed as a pragmatic way to help stakeholders with bounded rationality, and especially management, to cope with the enterprise entity they are concerned with. In a world shaped by the dynamics and complexity of reality, this accounting way constitutes the necessary blend of techniques of aggregation and dis-aggregation of information, taking into account costs of obtaining data, together with blending in consideration of law, custom of the society (professional duties included) and practices of the business world. Accounting constitutes an institutional way to cope with the firm as an enterprise entity.

Ultimately, equilibrium economics neglects the accounting system by placing the *price system* at the core of production and distribution. Equilibrium allows the market alone, seen as a system of prices in equilibrium, to govern the whole economic and monetary process. In order to surpass the limits of such a purely market economics, Coase (1992) invites us to pay attention to the actual institutional arrangements which govern the process of production (the firm) and exchanges (the market). In his opinion, inspiration for a new perspective is most likely to come through the stimulus provided by the patterns revealed by the systematic exploration of reality, particularly when the prime need is to break our existing habits of thought (Coase 1992: 718–19).

According to Coase, one of the highest tasks for the renewed theory may be to understand both the market and the firm by adding institutional economic substance into the usual "black boxes." The latter set the system of prices as the leading coordinating mechanism, but do not offer a satisfactory explanation of "the existence of management and of these apparently planned societies, firms, operating within our own economy" (Coase 1992: 715).

The institutional structure of production is at the core of "that little planned society" (Coase 1992: 716) that is the firm, and is also at the core of its special economic process as well. As Coase (1990: 13) explains, part of the "secret to the determination of the institutional structure of production [can be] found in the accounting system." The theory of the accounting system is hence a part of the theory of the firm. As Perroux had intended in the fifties, Coase aims to overcome a purely market view of the economic and monetary process. In order to understand the nature of the firm, Coase calls for proper consideration of the accounting system, and to relate it to the institutional structure of production.

The path appears open therefore to integrate business institutions other than the "property rights" in the theory and analysis of the firm as an institution and an organisation. Coase (1990) relies on the notion of opportunity cost to link management and accounting system within the firm, but he ignores the relevant implications of the financial accounting system that meets the business process of production with the (usually mandatory) institutional framework of representation and disclosure. Generally speaking, this framework is based on accounting principles, which, as Shubik (1993) recalls, the economic theorist must be able to fit into his theorising if he is to understand the world as it is. Therefore, both Coase and Shubik support the view that the accounting logic and structure are not merely irrelevant ephemera to the theorist. On the

contrary, their actual existence and peculiarities must be interpreted by a comprehensive theory capable of integrating them.

Drawn upon these insightful suggestions of Simon, Shubik and Coase, fostered by their realistic tenets, a new theoretical perspective on the firm finally implies a definite severance from the classic view of the lonely entrepreneur, proprietor and equity provider. Together with the price system, this figure prevents an improved understanding of the firm and of property (ownership and property rights) in the shifting context of the "twentieth century capitalist revolution." According to Berle (1965), the classical framework based on private property rights is factually unable to grasp the novelties of the firm and of its peculiar corporate form in the context provided by this capitalist revolution. The firm has acquired a definite dynamic and collective dimension that leads to and constitutes the field of overwhelming power. Setting aside merely legalistic reasoning, Berle delves into legal, economic and financial matters to understand this power. Drawing on the joint application of facts and insights, he seeks to interpret the firm as an active socio-economic institution and to understand its revolutionary economic, social, political and ethical implications.

As Weinstein claimed at the beginning of this part, the current state of the economic theory of the firm suggests the need for a new synthesis. All these insights, from different authors and different perspectives, open the way to a different view of the firm and of its institutional and organisational dimensions. The received view of the firm remains influenced by the neoclassical "black box" framed by the price system. In the "black box," no system of anything exists but prices in equilibrium. The firm is seen as a *nexus of prices*.

In order to overcome this bias, understanding the firm as a *nexus of contracts* (property rights included), be they incomplete, is not enough, since the firm may still be considered as merely a legal device which serves as a shallow nexus – deprived of any economic substance – for a mess of contracts among individuals or proprietors, whose only horizon is the quick pursuit of immediate wealth. This proprietary view maintains significant ties with the classical "invisible hand" perspective, and contrasts with relevant facts and insights about the dynamic and holistic characters that enter the drama of the firm under real dynamics and complexity.

Moreover, both legal and accounting logic (and principles) presently do not share this proprietary perspective on the firm and its role in economy and society. For instance, the practice of corporate donations, authorised by law and upheld by courts even without specific charter provision, is one where the contrast between the profit maximisation principle and actual reality is clear cut. Furthermore, concurrence is far from being as effective as a purely market framework pretends.

Recent crises in corporate governance and disclosure highlight the deficiencies and shortcomings of such an old-fashioned proprietary view of the firm. The managerial and enforcement devices it advocates imply poor policy and sometimes huge economic and societal losses. The third part of this volume seeks to overcome the proprietary bias by looking back to the time of high theory when these issues had already emerged and been discussed.

## Overview of the third part

A curious reader may be surprised by discovering all the apparently new questions that past scholars have already discussed at the time of high theory, i.e., from the beginning of twentieth century to the Second World War.

Canziani opens the third part by exploring the first emergence of some largely neglected theories of the firm that appeared in Continental Europe through the joint efforts of jurists, economists and accountants. In a sort of replication of Weigmann (1932), Canziani offers an instructive synthesis of this tradition based on three leading scholars, two from Germany (Schmalenbach and Nicklisch) and one from Italy (Zappa), well-known as the founders of Continental business economics of the firm. Through careful epistemological treatment, these authors reveal a self-contained theory prompted by relatively common questions and shared foundations. Their theory relies on a peculiar dynamic accounting view as the crucial clue for understanding the special economic and monetary process of production generated by the firm.

At around the same time, in the United States, the work of Littleton, who quotes Commons as a major reference of his thought, develops an institutional and dynamic accounting view for understanding the special economy of the firm. In the thirties, Littleton and his students explore legal and accounting matters in the attempt to bring business law, management and accounting closer together. "It should be possible [Littleton (1938: 84) writes] to visualize law and accounting as combining forces in setting standards which could aid management in giving well-balanced consideration to the interests of all parties concerned, as is fitting wherever fiduciary responsibilities obtain." Beyond doubt, Littleton was aware of the Continental development and was its key promoter in the American context, as proved by his references to European dynamic theorists, his activity as reviewer and his contribution to the German journal  promoted by Nicklisch.

The American and European developments on the matter thus appear to be related, as the legal-economic discourse also testifies. Indeed, the European institutional school of the enterprise entity ("*Unternehmen an sich*") fostered by Rathenau discussed legal-economic matters from a perspective that was very influential at that time. Berle and Means (1932), in the final chapter of this masterpiece, quote Rathenau as a key reference to support their own view of the firm (and of its peculiar corporate form) as a socio-economic institution. In this cross-national context, Kirat presents other scholars engaged in such important legal-economic developments. Delving into European and old American institutional economics, Kirat selects and summarises four scholars, two from France (Perroux and Ripert) and two from United States (J.M. Clark and Commons), all involved in theorising and analysing the firm and its relevant implications for economy and society. By distinguishing the firm and the corporation, Kirat explores each contribution from its legal-economic understanding of the capitalist enterprise coupled with the firm's economic and financial dynamics, the special income thus generated, and their accounting representation.

Avi-Yonah and Sivan complement the two first chapters with a historical perspective on the legal approach to the firm and to the peculiar legal form that

dominates the legal discourse, i.e., the corporation. This perspective reveals the relevant implications of theoretical views for the working of the legal system by analysing the transformations of judge and court decisions through time. Instead of odd dichotomies, Avi-Yonah and Sivan distinguish three peculiar views of the firm – either as an aggregate, an artificial person, or a real entity – that recurrently appear in different periods as perceptive standpoints. The real entity view is finally retained by virtue of its theoretical consistency, its heuristic and explanatory content, and its political and ethical implications for the social responsibility of firms.

The real entity view is fostered and further developed by Berle (1947) in a neglected paper on the legal-economic theory of the enterprise entity. This paper complements the critique of classical economic theory offered by Berle (1965) and explains why legal discourse has to leave the abstract world of legal forms to deal with the underlying socio-economic reality. A realistic and institutional understanding of this reality may provide an improved comprehension of the legal-economic substance of the firm seen as an enterprise entity. Dynamic and holistic characters factually enter the real drama of the firm, and key matters such as the group form and the accounting system may constitute the eligible field for improving current legal-economic wisdom. These matters are further developed in the fourth part of this volume.

Leaving the abstract world of legal forms and invisible hands further implies the severance from the self-assuring standpoint that "the [only] social responsibility of business is to increase its profits" (Friedman, 1970: 33). Only in the perfect world abstracted by a purely market perspective does the profit maximisation principle give clear-cut guidance for management and enforcement in setting priorities and deciding allocation of resources. However, in the real world laden with actual dynamics and complexity, such a clear-cut principle is proved to be unrealistic, unpractical, and ultimately immoral, as Anthony (1960) claims. In his clear and plain analysis, Anthony highlights the trouble implied by this preconception (of which the maximisation of wealth is simply a further refinement), and links the accounting role in the firm to the bounded rationality and institutional framework developed by Simon and others. Under real dynamics and complexity, the maximisation of returns can no longer provide a reliable guide for management and institutional enforcement. No *mechanism*, no "invisible hands," may ensure the attending of suitable social responsibilities and substitute the working together of management, organisation, and accounting. In this context, the accounting system factually provides a *satisficing* alternative to the logic of maximising profits or wealth, and constitutes an institutional way to represent and govern the firm and its peculiar economic and monetary process of production and allocation.

Adopting this integrated perspective, Stauss (1944) closes the third part of this volume with a synthesis that remarkably paves the way for further developments. This fully neglected institutional economist moves from the contribution of Paton and Littleton (1940) to an apposite criticism of the theories of the firm of his time, in particular those of Schumpeter and Knight, and expands upon the institutional notion of the "going concern" to understand the firm as a real entity having legal, accounting and economic features and implications.

## Overview of the fourth part

The fourth part of this volume discusses the contents and consequences of the theory of the firm as an entity by developing different fields of research. Instead of multiplying hypotheses, each paper purports to improve the comprehension of the matter by applying the framing idea of entity and its peculiar features in a consistent way, starting from general outlines with subsequent gradual filling in of details and implications. This way leads both to the enhancement of concepts and to the critique of received approaches, and would achieve a deepening of meanings together with a more comprehensive understanding.

As the previous parts suggest, an institutional way to understand the economic substance of the firm, i.e., to cope with the special economic and monetary process of production and allocation generated by the firm, is to delve into accounting and properly grasp its contribution to the theory of the firm as an entity. Biondi faces this challenge by developing the dynamic entity view supported by classic accounting principles. This view has two major implications. First, it implies increased recognition of the difference and articulation between the economic and monetary process generated by the whole firm and the process of exchanges related to external markets. Second, it enhances the institutional economic understanding of the separation between ownership, control and management beyond the irrevocably lost proprietary sovereignty.

This accounting contribution reinforces the interdisciplinary entity approach that integrates law and economics with accounting. Moreover, the notion of the firm as an entity has relevant implications for the legal-economic understanding of the firm as well. Law and accounting are the twin pillars that sustain a new positive assessment and provide new theoretical insights for improving the theory and analysis of the firm.

Both Gindis and Manfrin deal with the "legal-economic system" involved by the firm and try to distinguish its economic substance as an enterprise entity from the legal forms involved by its actual dynamics. This distinction implies a movement from the enterprise taking just one legal form towards the overarching enterprise entity that involves a web of actual institutional arrangements, including the networking of corporations as a group, labour contracts, security exchanges, fiscal and monetary regulations, and so on. By starting from this distinction, the usual legal-economic approaches that identify the firm with the corporation lack most of their theoretical and heuristic force, and the institutional legal-economic perspective can be improved and supported.

In particular, Gindis disentangles the ethical-political view from the ontological view on the legal-economic framework of the firm. Even though both perspectives can be perceptive, Gindis chooses to point to some ontological building blocks to properly account for the firm as a real entity. These building blocks overcome the deficiencies of aggregate and fiction views, as well as of purely normative theories that misunderstand the reality of the firm and might imply unexpected and counter-productive consequences when followed as pragmatic guidance – even though someone may scrutinise the donkey theory "as if" it flew,

this does not appear the best way for reasonable implications and recommendations. On the contrary, among the dynamic realism building blocks, Gindis stresses three features of the firm early recognised by Freund (1897), namely its unity, its distinctiveness and its identity in succession. The firm then is not reducible to its parts or members, and is certainly not a fiction, legal or otherwise, but a real entity that exhibits a relative degree of cohesion and durability through time. Gindis applies this view to the logical and methodological foundations of the theory of the firm, and effectively sketches its consequences on the overwhelming practical problems of enterprise group regulation and network organization.

Manfrin brings such a realistic undertaking further by exploring how regulation affects the accomplishment of the economic activity of the firm. Starting from the insightful distinction between economic substance and legal forms, Manfrin considers the regulatory framework as a legal-economic system driven and shaped by both management and polity (either government policies, or legislation and courts). To corroborate this standpoint, Manfrin deals with both theories and practices. Hence, he shows the relevance of the real entity view of the firm contrasted with three other legal-economic approaches to the firm, namely as an artificial person, as a nexus of contracts or one of property rights. Manfrin then analyses the relevant implications of the real entity view for the legal-economic understanding of the enterprise group and the question of the common interest.

Marzo deals with the economics and finance of the firm as an entity. He provides an impressive and critical analysis of the basic premises of the neoclassical financial theory, and highlights the numerous neglected or misunderstood matters that a real entity view of enterprise finance can better explain. The link between accounting and enterprise finance is further developed from the entity viewpoint, expanding upon Simon and Anthony. This theoretical perspective provides a heuristic model of the *process of financing* the enterprise entity by recourse to featured financial sources subject to real dynamics and complexity.

Enterprise governance can be seen as the pragmatic synthesis of accounting, legal-economic and financial issues. In conclusion, as a sort of summary of the fourth part, Moore and Rebérioux deal further with the matter of governing the firm as an enterprise entity. Taking some recent events as reference point, Moore and Rebérioux's positive and normative assessments question the current wisdom about the divergence of interests between managers and shareholders by stressing the overwhelming evidence of the actual controlling alliance between leading managers and influential shareholders. This provides an original rationale for the recent crises in corporate governance and disclosure. To improve on their analysis, Moore and Rebérioux revisit the very contribution of Berle and Means on the firm as a socio-economic institution and trace its relevant implications for current theoretical and applied debates.

In summary, the authors of this book view and analyse the firm by the interdisciplinary approach of the enterprise entity. They understand this entity as a whole and a dynamic system. Therefore, they pay careful attention to the institutional economic substance of its inner organisation, especially legal-economic, accounting and financial systems. These are related to the nature of basic

learning, diffusion, and information processes, to the closeness of products, technologies, resources and internal organization, to the working framework of rules and norms, and to consideration of business strategy, business finance, sociology and regulation. In so doing, the aim is to improve the current understanding of the firm as an institution and an organisation concerned with the actual dynamics and complexity of reality.

## Acknowledgment

As convener of the book project, I would like to acknowledge the valuable commitment by Rob Langham to this collection of essays throughout its publication process.

## References

Anthony, N. Robert (1960), "The Trouble with Profit Maximization," *Harvard Business Review*, 38 (6), November–December: 127–34, reprinted in this volume.

Berle, Adolf A. Jr (1947) "The Theory of enterprise entity," *Columbia Law Review*, 47 (3) April: 343–58, reprinted in this volume.

—— (1965), "The Impact of the Corporation on Classical Economic Theory," *Quarterly Journal of Economics*, 79: 25–40, reprinted in this volume.

Berle, Adolf A. Jr. and Means, Gardiner C. (1932), *The modern corporation and private property*, New York: Harcourt, Brace & World.

Coase, H. Ronald (1990), "Accounting and the theory of the firm," *Journal of Accounting and Economics*, 12, 1990: 3–13, reprinted in this volume.

—— (1992), "The Institutional Structure of Production," *American Economic Review*, 88 (4), 1992: 713–19.

Durkheim, Émile (1930), *De la division du travail social*, Paris: PUF, 2004.

Freund, Ernst (1897), *The Legal Nature of Corporations*, Chicago: Chicago University Press.

Friedman, Milton (1970), "The social responsibility of business is to increase its profits," *New York Times Magazine*, September 13: 30.

Littleton, A. Charles (1938), "A substitute for stated capital," *Harvard Business Review*, 17 (1), Autumn: 75–84.

Paton, A. William and Littleton, A. Charles (1940), *An Introduction to Corporate Accounting Standards*, USA: American Accounting Association.

Rathenau, Walter (1918), *Von kommenden Dingen*, Berlin, trans. E. and C. Paul, "In Days to Come," London, 1921.

Shubik, Martin (1993), "Accounting and its relationship to general equilibrium theory," in M. Shubik (ed.), Accounting and Economics, *Economic Notes* by Monte dei Paschi di Siena, 22 (2): 226–34, reprinted in this volume.

Simon, A. Herbert (1979), "Rational Decision Making in Business Organizations," *American Economic Review*, 69 (4), September: 493–513.

—— (1991), "Organizations and Markets," *Journal of Economic Perspectives*, 5 (2), Spring: 25–44, reprinted in this volume.

Stauss, James H. (1944), "The Entrepreneur: The Firm," *Journal of Political Economy*, 52 (2), June: 112–27, reprinted in this volume.

Weigmann, W. (1932), "Legal and Economic Concepts of the Balance Sheet in Germany," *The Acconting Review*, 7 (2), June: 103–106.

# Part II

# On the economic theory of the firm as an institution and an organisation

# 3 The current state of the economic theory of the firm

## Contractual, competence-based, and beyond

*Olivier Weinstein*

## Introduction

Together with the market – and the modern State – the firm is one of the main institutions of capitalism. We can even consider it the central institution, despite the fact that contemporary economic discourse tends to attach more importance to the market. As Simon points out so pertinently (1997: 35):[1]

> . . . Any creature floating down to our Earth from Mars would perceive the developed regions to be covered mostly by firms, these firms connected by a network of communications and transactions that we know as markets. But the firms would be much more salient than the markets, sometimes growing, sometimes shrinking, sometimes dividing or even swallowing one another. Surely they would appear to be the active elements in the scene.

A view shared by Chandler (1992: 79):

> Although this is rarely mentioned in the literature, "for profit" firms have been and still are the primary instrument in capitalist economies for the production and distribution of current goods and services and for the planning and allocation for future production and distribution.

The firm as it appears in modern economies is also a complex and coherent structure, founded on organisational architectures of extremely diverse forms (Aoki 2001). Thus, the modern company can adopt different configurations, depending on the country and the activity; it has undergone profound transformations throughout the history of capitalism, described particularly clearly in the works of Chandler, and it is currently experiencing major changes. It is not surprising, under these circumstances, that attempts to understand the "nature" and characteristics of the firm and the attributes of the modern corporation raise a large number of questions. The firm resists all attempts to reduce it to a unique function or a simple principle. And this is exactly what the current state of economic literature on the firm demonstrates.

Over the last thirty years or so, the theory of the firm has been reconstructed on the basis of the issues raised by Coase (1937). To begin with, this steered the analysis in a particular direction, by focusing on the question of the modes of coordination between individual economic agents at the most general level, on the basis of an opposition between the market (price coordination) and the firm (hierarchical or "administrative" coordination). Within this context, the firm is defined as a specific mode of coordination, alternative to the market, and as a response to market failure. This raised two questions: why and under what conditions is a different form of coordination required, other than that of the market? And what are the fundamental characteristics of coordination by the firm? What is the "nature" of the firm? Subsequently, besides the calling into question by certain observers of the opposition between firm and market, the analyses and debates have been extended to include other questions, particularly concerning the internal structure of the firm, forms of ownership, control and governance, and analysis of the dynamics of firms and organizations. This extension of the investigation – understandable given the complexity of the subject – has gone hand in hand with theoretical questioning of the foundations of the analysis of organisations, or even institutions, and, more fundamentally, of the foundations of micro-economic theory itself.

This chapter is not intended to present an exhaustive overview of the contributions of the different works on the theory of the firm. Our aim is simply to highlight certain lessons and questions which stand out in the literature, concerning our understanding of the firm, its past and present forms and the possible principles of an economic theorisation of the firm. We shall do so from the perspective adopted elsewhere in this book: by seeking to understand how the firm can be apprehended as a specific entity, situated at the heart of a given economic system, as described by Chandler and Simon, after many other economists and historians.

The first question we must tackle concerns the precise object of the theory of the firm: as Hodgson pointed out (1999), Coase – and many subsequent works on the subject – remains very vague about what he actually means by the "firm," a term which seems to cover any organisation whose aim is the production of goods or services. Taking a broad definition of the concept of service, this could include just about every organisation. But the implicit focus of Coase's enquiry is the capitalist company: – in other words a specific form of organisation, situated historically, which cannot be reduced to the simple implementation of a hierarchical principle, which can itself take a wide variety of forms. Indeed, although other possible forms of company are occasionally evoked, the capitalist firm (and more particularly the "corporation") is, whether explicitly or not, the essential object of the greater part of the theory of the firm. And this, the "primary instrument in capitalist economies for the production and distribution of current goods and services," will be the centre of our attention here. This leads us to focus on the relations between three categories of actors: the directors (or managers), the employees and the shareholders. In our opinion, most of the current theory of the firm revolves around these relations and the problems they raise.[2]

This leads us to introduce another major reference, which, alongside Coase, is essential for understanding the debates about the firm. This is Berle and Means. Berle's and Means' analyses on the development of the stock company and the question of its *ownership* and *control* provide a vital complementary insight, because of what they say about the development of capitalist society and private property. In particular, the works of Berle and Means are important because they propose a method of theorization of the *capitalist* firm, marked by three main features: (i) The characteristics, functioning and behaviour of the firm are analysed in terms of the relations between the different groups, *with their own specific interests*, which constitute the firm and maintain close relations with it: the shareholders and managers, but also the employees and the suppliers of credit; (ii) the central issue is then to determine *who controls* the firm and what the relations are between ownership, management and control; (iii) the institutional framework, in this case the stock company, the legal principles on which it is founded and the development of capital markets, play an essential role in structuring the firm in particular, and society in general. The works of Berle and Means also demonstrated a clear distinction between *firm* and "*corporation*," a distinction which is rarely made.[3] So, if the questions originally raised by Coase (1937) mainly concern the specificity of the "administrative" coordination inside the firm, and hence the nature of the relations between employer and employees, Berle and Means lead us to the question of the status of managers and the problems posed by the "nature" of the public corporation.

Theorisation of the firm is mainly built around the issues raised by Berle and Means, by Coase and by later "competence-based" approaches. To give an account of these questions, we shall proceed as follows: the first two sections will be devoted to what remains today the dominant approach, within the framework of which the major questions were first raised, i.e. the neo-institutional and contractual movement. The first section will deal with the questionings of Coase and labour relations, which will lead us, as we shall see, to the question of the role of property; the second section, starting with Berle and Means, will lead us on to agency theory and the questions it raises. The third section will explore analyses which consider the firm chiefly as a collective space for the organisation of production, which will lead us to "competence-based" approaches and to the perspectives opened up by Herbert Simon and the evolutionists. To conclude, we seek to demonstrate why, as we see it, only the specifically institutionalist approaches can provide a solid basis for understanding the nature and characteristics of the capitalist firm and its forms and transformations.

## The Coase legacy: the nature of the firm, employment relation and property

Nowadays, it is customary to consider the economic theory of the firm as being divided by a primary opposition between two main approaches: the neo-institutional or contractual perspective and the "competence-based" or "knowledge-based" perspective.

The former, which dominates the current economic analysis of firms and organisations, derives more directly from Coase's questionings. Although there are contrasting, if not conflicting theories within this approach, there is profound unity concerning the nature of economic (and social) relations between actors, conceived as free contractual relations. This means that all institutions or organisations are viewed as the result of agreements between individuals, the key questions being the problems of the definition (the "design" of the contract), and "enforcement" of such agreements, and the costs resulting from these difficulties, which are referred to as "transaction costs" or "agency costs." Within this contractual perspective, whose origin can be traced back to the early development of liberal European political philosophy, present neo-institutional economics adds two crucial – and questionable – dimensions: (i) the problems of coordination between people are held to result mainly, if not exclusively, from imperfections of information and information asymmetries between parties, and (ii) the central purpose of the economic theory of the firm (and more generally of organisations) is to identify the most efficient forms of organisation in different contexts, forms which are supposed to emerge from the free interactions between agents. As clearly expressed by Masten (2002):

> In its most abstract form, the positive theory of the firm is simply an extension of standard rational-choice principles: Given two or more alternative ways of governing their relations, transactors will choose to organize "internally," say, *Gf*, rather than "in the market," *Ga*, when they expect internal organisation to yield a higher joint surplus than the alternative(s), $Vf > Va$. A theory of the firm adds content to this framework by providing guidance as to the factors that influence the outcome of that comparison.

In this approach, the firm is studied as a specific contractual form of coordination. The firm is viewed as *a nexus of contracts* between *individuals*. However, most analyses lead to consideration of the relations between *groups* (managers, shareholders, workers . . .), a point which clearly has very important implications for the definition and characterisation of the firm. The differences between various contractual approaches (and other approaches as well) may result from the way in which behaviour and interactions are analysed, but they may also result from the definition of the key groups and the relative importance attached to different relations. In other words, a theory of the firm must involve a specification of the "structure," or the architecture of the firm's organisation, and of its institutional foundations, including the formation of social groups.

The literature generally considers that contractual and neo-institutionalist works propound three main theories of the firm: agency theory, transaction costs theory (TCT) and the theory of incomplete contracts and property rights. Like Gibbons (2005), we believe that transaction costs theory in fact comprises two theories. The better-known of these two is called the "hold-up" theory; Gibbons refers to the other, which is markedly different, as the "adaptation" theory. To our mind, the latter represents a more radical departure from the neoclassical

postulates which form the basis of both agency theory and property rights theory. To this we can add, as Blair and Stout (1999) did, what has been called a "team production approach," focused on the problems of coordination between the members of production teams. This last approach is present, as we know, in the pioneering article by Alchian and Demsetz, but it has a more fundamental dimension that we shall tackle in the third section, involving the taking into account of production problems *per se*. In many respects, these different approaches don't answer exactly the same questions, and they propose different conceptions of the firm. This is the point we wish to underline here. To do so, it may be useful to start with the issues raised by Coase on the one hand, and by Berle and Means on the other. As Bolton and Scharstein (1998) observed, full understanding of the firm requires the integration of both of these points of view – a project that has, for the most part, yet to be realised.

As we know, Coase took as his starting point an opposition between firm and market, which logically implies that the firm is considered as a specific entity. If we accept this point, two questions arise: (i) What is it that fundamentally characterises the firm (what is the "nature of the firm"?) or, to put it another way, what is it that differentiates coordination in the firm from coordination through the market? (ii) Why are certain activities performed within the firm, rather than being coordinated by the market, and what determines the choices and limits of internalisation?

In most cases, the first question is dealt with by characterising coordination within the firm as being coordination through hierarchy and through the exercise of the power of authority. Hierarchy and authority are usually treated as synonyms, or at least as going hand in hand. A distinction must be made between the principle of authority that is at play within the firm – at different levels and in different forms – and the hierarchy that is expressed in the organisational structure of the firm, as described by Chandler, for example. We shall come back to these points later. It is also important to consider what exactly is meant by the "integration" of an activity, and how this may change its mode of functioning. The standard reply, which often remains implicit, is that integration means "the unification of control rights" (Gibbons 2005). This in turn leads the question of what we mean by control, who has the power of control (and on what basis), what is control and how is it exercised? In other words, we must investigate the nature and contents of the power or powers at work within the firm. The characterisation of the firm in terms of hierarchy also calls for the analysis to be focused on the nature of the employment relationship.

It is the second question that has received the most attention, to the point where the "make or buy" problem often appears as the main – if not the only – subject of the economic theory of the firm. The best-known reply to this is given by the theory of transaction costs, specifically the version known as the "hold-up" theory (see Williamson 1979, 1985 and Klein, Crawford and Alchian 1978): internalisation is a response to the problems of conflict over the appropriation of the quasi-rent produced by a transaction, in the case where this transaction

involves important specific assets. Coordination through authority (fiat) can then achieve lower transaction costs than haggling.

In our opinion, the reply provided by the theory of incomplete contracts developed by Grossman, Hart and Moore is fundamentally no different, although it proposes a different type of theorisation, especially with regard to the cause and implications of the incompleteness of contracts.[4] The incompleteness of contracts clearly occupies a central role in the argumentations of contractual approaches, and yet it often remains a poorly defined concept.[5] In particular, it is questionable whether it has any meaning beyond the hypothesis of limited rationality rejected by Grossman, Hart and Moore. The theorisation of these authors appears, first and foremost, as an attempt at rigorous formalisation of the TCT analysis of integration, making it possible, in addition, to take into account both the benefits *and the costs* of the integration, where TCT presents two different types of explanation, by bringing forward the problems of bureaucracy, and above all the question of incentive and the impossibility of "selective intervention," to explain the limits of integration (Williamson 1985: ch. 6).

However, the search for a rigorous formulation requires the theory of incomplete contracts and property rights to be more precise about the identity of the firm and about the question of control and the relationship between control and ownership (and between ownership and incentive). First, the firm is defined as a set of (non-human) assets. It is presented as "a bundle of assets under common ownership (and therefore, common control)" (Blair and Stout 1999: 260). Here, however, two important points should be noted: (i) control and property are assimilated, and the definition of control is linked to the incompleteness of contracts; (ii) as the passages quoted in the footnotes show,[6] there is a certain vagueness about who is the actual owner of the assets: the company or an individual (who is, one would imagine, by this very fact, the owner of the company[7]). Like Holmstrom (1999), we believe that the formalisation proposed actually means that it is an individual, and not the firm, who is considered to possess the assets and to enter into contracts with the other parties. This view, which operates within a logic of inter-individual contracts, makes it difficult, if not impossible, to understand the firm as a coherent whole, and to take production, in the strict sense of the term, into account.

This places the ownership of assets at the heart of the analysis, but with a rather special conception of ownership, far removed from the juridical view.[8] The distribution of property is shaped by the level of specific investments made at the outset, and this will determine which distribution (separation of assets between two agents, or unification under the control of one of them) has the optimal outcome. Ownership derives its importance directly from the incompleteness of contracts. When it is impossible to specify the use of an asset in every possible situation in advance, by contract, it is ownership which determines who will have the power to decide. The ownership of an asset is essentially defined as the possession of a "residual control right," i.e. the right to decide the uses of an asset, within the limits of what is allowed by law, custom or

previous contracts, and therefore the right to exclude other uses of the asset. This is why Hart (1995: 29) wrote: "ownership is a source of power, when contracts are incomplete." Further, this power over things extends into power over workers, to the extent that these latter possess asset-specific skills, i.e. skills which they can only exercise when given access to certain assets: "this authority over assets translates into authority over people" (Hart and Moore 1990). It is hardly surprising that Hart (1995: 5) establishes a link between this conception and Marx's analysis of the relationship between capitalists and workers, and the importance attached to the ownership of the means of production.

Yet we should point out that in these analyses, this control (or authority) over the employees has very limited contents: it boils down to the power to exclude workers from use of the asset, in other words the power to lay off.[9] But this power exists just as much in a contract for the provision of services. Thus, the employment relationship in property rights theory cannot really be distinguished from the service relationship, in other words a market relationship. To really distinguish between firm and market, a clear distinction needs to be made between sales contract and employment contract,[10] and between inter-individual contract and contract with a company.[11] This can only be achieved by taking into consideration the more global power of the employer to decide and control the employees' actions. This is what authors, like Simon (1951), for example, generally mean when they refer to hierarchy. One of the reasons why this aspect is overlooked is because the question of the organisation of production, in the strict sense of the term, and of the coordination *between workers*, is not studied (a point recognised by Hart and Moore 1990: 1152). Another aspect is also neglected: the employer, the owner of the assets, is also the owner of the product, and this can be considered one of his main characteristics (Ellerman 1992).

Lastly, property rights theory assumes that residual control rights are non-transferable, from which it ensues that we are dealing with a firm without managers (Holmstrom 1999; Gibbons 2005). We are therefore tempted to say that property rights theory overlooks not only the employment relationship but also the relationship between shareholders and managers.

Williamson puts forward another justification for integration. This makes it easier to understand what really opposes transaction costs theory and property rights theory: the former is founded on Simon's vision of behaviour, the hypothesis of limited rationality, and focuses on the *course of contractual relationships over time*. Williamson, indeed, puts great stress on this point: "process matters" (Williamson 1991: 98). This aspect is derived directly from the hypothesis of limited rationality and its consequence, the incompleteness of contracts: the major problem raised by a long-term relationship between two interdependent parties is the management of unforeseen events, which cannot be anticipated and cannot, therefore, be taken into account in the formal contract. Modes of governance are differentiated from each other precisely by their manner of dealing with these unforeseen events, the specificity of integration being that it gives one of the parties the discretionary power to take decisions in response to

events, and so to enable adaptation of the organisation *without renegotiation*. So the specificity of integration is that it increases the adaptivity of the organisation to its environment, and thus, in a way, favours *dynamic efficiency*. The point of departure of this analysis does lie within the hypothesis of limited rationality, which requires that: "Exchange will be facilitated by modes that support adaptive, sequential decision-making" (Williamson 1991: 93).

Williamson's conception involves a broader vision of the authority specific to the firm than that to be found in the concept of "residual control rights," although it would not be unfair to say that the contents of the authority relationship remain largely undefined. Here, to gain a better understanding of the view of the firm implicitly present, it may be useful to explore its links with the representation of the employment relationship proposed in Simon's seminal article (1951).[12] The aim of this text was to define precisely the contents of the "authority relationship" specific to the employment contract, in contrast to an ordinary contract or sales contract. Designating the employer as B, the employee as W and the set of tasks that W can accomplish as x:

> *B* exercises an authority over *W* if *W* permits *B* to select *x*. That is, *W* accepts authority when his behaviour is determined by *B*'s decision. In general, *W* will accept authority only if $x_0$, the *x* chosen by *B*, is restricted to some given subset (W's "area of acceptance") of the possible values.
>
> Simon (1951: 294)

Further on, Simon (1951: 305) adds: "in an employment contract, certain aspects of the worker's behaviour are stipulated in the contract terms, certain other aspects are placed within the authority of the employer, and still other aspects are left to the worker's choice."

Simon's analysis is important because it affirms the specificity of the employment contract, an essential element for grasping the characteristics of the modern firm, while managing to give a very general characterisation of this specificity and while remaining at the level of inter-individual relations. The analysis would benefit from more precision about how the distinctions are made between the three domains evoked by Simon: that which comes within the scope of the actual contract, that which involves the employer's authority, and that which involves the worker's autonomy. This could then be taken as the basis for an exploration of both the fundamental characteristics of the capitalist firm and the various more specific forms it can take. To go forward in this direction, the different forms of the division of labour and the characteristics of the organisation of productive work in the firm must be examined in more detail, and this, as we shall see, means looking beyond the contractual dimension.

We also need to investigate the foundations of this employment relationship: on what grounds does the management of a firm dispose of the powers conferred on it? To what extent and for what reason do the workers accept this authority? The reply favoured by all the contractual approaches appears to be that the features of the employment relationship, whatever they may be, are the

result of free agreement between the parties, the employees accepting this type of contract to the extent that they expect to benefit from it. We can add to this, as property rights theory does, that it is the possession of productive assets by others that leads – or even obliges – certain workers to accept this type of relationship. This introduces a relation between *property regime* and form of organisation, an essential aspect for any institutionalist approach, but which calls for a more thorough exploration of what exactly this property regime is.[13]

There remains the fact that we can also consider that the effective authority of the directors (or managers) over the employees isn't necessarily based on ownership, or at least that ownership, although it confers formal authority, is not sufficient to guarantee *real authority* (Aghion and Tirole 1997). This is another reason why full understanding of the employment relationship – and of the firm – cannot be achieved solely on the basis of consideration of legal and explicit contractual forms. This is demonstrated, in a certain fashion, by works which focus on the role of "implicit" or "relational" contracts (see, in particular, Baker, Gibbons and Murphy 2002). And, as Zingales (2000: 1633) observed, the contractual theory of the firm is radically modified when we take implicit contracts into account, to the extent that this means that "a firm is not simply the sum of components readily available on the market but rather is a unique combination, which can be worth more or less than the sum of its parts."

This doesn't mean that ownership – and the legal system in general – is without importance. The legal system plays a major role as the foundation and "legitimation" of the employment relationship, by affirming the difference between commercial contract and employment contract, especially in terms of the definition of the employee's obligations (Masten 1988). However, as Holmstrom and Roberts (1998) pointed out, the incentive to make specific investments – or to perform the right task – can be provided in several ways, not simply through ownership, and the property system has multiple functions, far beyond what is taken into consideration here.

We can make one last point on this subject: when ownership is taken into account it is obviously necessary to specify who the owner is and what he owns. Now, property rights theory concerns the possession of assets *by individuals*, not by the firm. This is what Holmstrom (1999) criticised when he observed pertinently that one of the essential characteristics of the modern firm is the fact that the firm itself owns most of the productive assets. To understand the modern firm, we must start by recognising the fact that it constitutes a specific entity in its own right. It is very much the firm which owns the assets, as it is the firm – and not the directors or managers – which enters into contractual relationships with the employees and other parties. As Chandler (1992) said: "the firm is a legal entity – one that signs contracts with its suppliers, distributors, employees and often customers." As an entity, it is also, in itself, "a tradable commodity" (Putterman and Kroszner 1996), and this is a far from negligible characteristic.

This leads to a completely different view of the firm. Consequently, as Holmstrom (1999) pointed out, one of the most important issues is to explain

why the quasi-totality of productive assets is owned *by firms*, with most employees possessing no more than their skills. This means that we must start with the constituent attributes of the capitalist firm, before moving on to grasp the contractual – and legal – forms that it mobilises.

In their work, Rajan and Zingales propose a significant new departure, concerning the question of the nature of the firm and the relations between property and power. Their analysis still takes Coase as the starting point, by taking the authority relationship as the primary characteristic of the firm. The analysis is then built on an investigation of the foundations of this power. From property rights theory, Rajan and Zingales (1998) adopt both the characterisation of power in the firm as a residual (and non-contractible) control right and the importance of this control in "protecting" specific investments, but they reject the postulated exclusive link between control and the ownership of physical assets.

The starting idea is that the entrepreneur – or the manager – can acquire and maintain power over the workers, in other words over the human capital, to the extent that he controls, through ownership *or by other means*, a unique "critical resource," which could be a material asset, but could equally well be certain people or knowledge. The main problem arises when this critical resource is not an alienable asset that can be controlled by legal measures and allocated by the market. What matters is the *access* to critical resources, in other words "the ability to use, or work with, a critical resource" (Rajan and Zingales 1998: 388), which is quite different from possession. As with Hart, the complementarity between resources is an essential instrument in the control of non-appropriable assets, especially human capital: "By controlling a critical resource, an entrepreneur can influence the accumulation of specific investments so as to build complementarities between the person the entrepreneur seeks to have power over and her critical resource" (Zingales 2000: 1645).

This aspect is all the more important for the authors because it forms part of an overall vision of transformations in the firm. According to this vision, now widely adopted, critical resources are more and more often constituted of human capital, rather than of material assets. This result in a certain conception of the identity of the firm: *a firm is more than asimple collection of assets*, particularly because *the employees belong to the firm*. "When firms' assets represented the main source of value and control, it was perfectly legitimate to approximate a firm with its physical assets." (Zingales 2000: 1643); this can no longer be the case when intellectual capital becomes one of the principal sources of value, and when the firm can no longer be indirectly controlled solely through the possession of physical assets.[14] Thus, it is a "unique combination of assets and individuals that constitutes a firm" (Zingales 2000: 1626).[15]

This also leads Rajan and Zingales to criticise the view of the firm as a nexus of *explicit* contracts, preferring a vision which takes into account "a nexus of explicit and implicit contracts." As we have said, this constitutes a major break, in that it leads to the firm being considered as a whole, built around an "organizational capital," which "cannot be reproduced instantaneously, through legal

arrangements" (Rajan and Zingales 2001b: 842). This characterisation of the firm is similar to that found in competence-based theories, to which the authors refer repeatedly, but with an additional central element: they take into account questions of power and the control of knowledge and skills,[16] the complex relations between conflict and cooperation within the firm (Rajan and Zingales 2000), and the question of the sharing out of the surplus created by the firm. All this involves consideration of the role of *internal organisation* in increasing the value of the firm. The essence of internal organisation resides in the different access that agents have to the critical resources which make up the heart of the firm, and in the way in which property rights and access capacities are combined. Another implication is that the *economic* definition of the firm no longer strictly corresponds with its *legal* definition. This is connected with the fact that the growing importance of human assets – which are inalienable – has led to a *separation between ownership and control* (Rajan and Zingales 2001b: 841).[17] This in turn calls into question the role of property: the ownership of assets by an agent can reduce his incentive to make specific investments (Rajan and Zingales 1998: 406). This is what the authors describe as the "*dark side of ownership*" (Rajan and Zingales 1998: 390).[18]

The programme of Rajan and Zingales is particularly ambitious. They combine a contractual approach with competence-based dimensions and with an analysis of the historical transformations of capitalism, without forgetting questions of corporate finance – which we shall not go into here. Let us focus on one question: to what extent do Rajan and Zingales break with the contractual view of the firm, and do they achieve a real characterisation of the firm as entity? Their approach remains within a contractual analytical framework, and yet it calls the contractual vision into question. In their emphasis on the (growing) incompleteness of contracts, and even more on the importance of implicit "contracts," it is clear that the scope and importance of explicit contracts in the analysis of the firm have been diminished: most of the relations and actions around which the firm is structured occur *out of contract*. So the contractual vision is gradually dissolving. The authors' analysis also shows the absolute necessity of giving a central position to the analysis of production and internal organisation, even if we must look elsewhere to find the essential elements for this and even if the connections between the different dimensions have, for the most part, yet to be constructed.

However, this doesn't mean that we should ignore the importance of the contractual – and legal – dimension of the firm. It is just that the contractual perspective cannot account for all the attributes of the firm. It may even appear to be inconsistent with some of its essential features, particularly when it comes to grasping the firm as an entity endowed with its own specific and long-lasting existence, which cannot be reduced to the system of contracts existing between given agents at any given moment. The only alternative would be dramatically to extend the concept of contractual system – in the sense of an institutional system – by defining a whole set of positions, and thus social groups, which exist independently of the individuals who occupy them.

We must also recognise that the contractual dimension is more than just a theoretical choice; it also reflects the nature of the modern social, economic and political system which tends to treat the contract as the fundamental form of relationship between agents. The opposition between contractual relationship (reversible and transitory) and the firm (as long-lasting entity) is a constituent part of our economic system, especially nowadays. Indeed, the organisation of the firm – in its different dimensions, and particularly in terms of finance (and ownership) and in terms of the organisation of production – is an attempt to manage this very tension: how to ensure both the continuity of the firm's activity and its flexibility (ability to enter and leave an activity, to redeploy its resources and so on). This is true both on the financial level (in the contradictory relation between the liquidity of financial assets and control), and on the productive level (through sub-contracting, networking, labour force mobility, etc.). This goes some way to explaining the growing distance between the legal and economic delimitations of the firm.

Essentially, the questions raised by Coase concern the conditions of the choice between firm and market, and hierarchy and the authority relationship as the distinctive features of the firm. What remains to be taken into account is the fact that the main institution of capitalism over the last century has not been the individual firm so much as the corporation. As Blair and Stout (1999) remark, property rights theory is not a theory of the corporation, and it remains a theory without managers (Holmstrom 1999; Gibbons 2005). To tackle these questions, we need to consider the possibility for the owner of the assets to delegate his decision-making powers. This in turn leads to the question of the relations between shareholders and managers, to the questions raised by Berle and Means, and to agency theory.

## The (public) corporation and the relations between shareholders and managers

Agency theory remains a key reference in today's dominant economic view of the firm, all the more so since the question of corporate governance has come to the fore. We feel that this theory is best understood in the light of the questions raised by Berle and Means – it can be seen as a reply to the theses put forward by these authors. In many respects, it is also a reply to those who, following Coase, present the firm as a response to market failures, and as an organisational form distinct from, or even opposite to, that of the market.

Beyond their observation of the separation between ownership and control of the firm, Berle and Means considered that the developments they were witnessing necessitated a profound rethinking of the conceptions of firm and ownership. The firm should be considered as an *institution*[19] and a "*social organization*", which:

> . . . involves the interrelation of a wide diversity of economics interests, – those of the "owners" who supplied capital, those of the workers who

> "create," those of the consumers who give value to the products of enterprise, and above all those of the control who wield power.
>
> (Berle and Means 1932: 310)

The firm exists in its own right, as a specific entity, beyond the changing personalities of its shareholders, workers and managers.[20] And this firm can be regarded both as a confrontation between diverse interests and as a "concentration of powers." This raises the question of how this power can be regulated and how protection of and arbitration between the different interests can be achieved.

For Berle and Means, this meant calling into question the "traditional logic of property," according to which "by tradition, a corporation 'belongs' to its shareholders . . . and theirs is the only interest to be recognized as the object of corporate activity" (Berle and Means 1932: 310). For the authors, this "doctrine of strict property rights" cannot be applied to the *passive* ownership by shareholders of large corporations. The latter, by renouncing control and the responsibility for management, have thus renounced "on a quasi-contractual basis" the right to have the firm managed solely in their interests (Berle and Means 1932: 311–12). To this must be added the importance of the liquidity provided by the development of capital markets which, as Aglietta and Reberioux (2004) have shown, "assumes *precisely* a total detachment between the person and the property." The development of the large corporation and of modern capitalism does indeed signify, as Rathenau foresaw, the "depersonalization of ownership."

Agency theory takes the exact opposite view to these theses. The delegation of decision-making powers, at many different levels, is one of the major characteristics of the functioning of both the firm and the wider economy. This delegation is justified by "organizational complexity" and the dispersion of knowledge between different individuals, especially "specific knowledge,"[21] which explains the gains resulting from the division of labour (particularly in the realm of management) that enable decision-making powers to be conferred on those who possess the appropriate specific knowledge.

This can lead to a conflict between the allocation of ownership and the allocation of decision-making powers. The capitalist system has solved this dilemma "by granting alienability of decision rights to decision agents" (Jensen and Meckling 1992–98: 103).[22] It is on the strength of this point that agency theorists reaffirm a principle of "strict property rights." To do so, they can draw on the economic theory of property rights first developed in the 1960s, by writers such as Alchian, Demsetz and Barzel.[23] The aim of these authors was to strengthen property rights theory, so as to make it one of the foundations of all our society's organisations and institutions. They laid particular emphasis on the complexity of property rights systems, which are constructed on the premise that the rights are *alienable, partitionable and separable.* Thus, a property right is not simply an individual's right to decide on the uses to which a good is put, but also the right to transfer this right to another person, and the possibility of

sharing it out among several individuals. This laid the foundations, in particular, for the public corporation, in which the shareholders, who share ownership of the capital, delegate to managers the right to control the use of that capital. And it also formed the – admittedly fragile – basis for the premise that managers should be considered as no more than the agents of the shareholder-owners, their relationship being treated as a specific example of an agency relationship. The problem is then to take into account all the measures of incentive and control required to ensure that the managers (the agents) behave in accordance with the interests of the shareholders (the principals), leading, as we know, to the analysis of agency costs.

This goes hand in hand with the view of the firm that clearly rejects both the idea of the firm as a specific entity and the opposition of the firm and market as radically different forms of coordination. For Jensen and Meckling (1976), organisations, particularly the firm, are no more than "legal fictions which serve as a nexus for a set of contracting relationships among individuals." This is an important lie in the characteristics of the different contractual relationships, and contractual relationships *between individuals*. This general vision of the firm as a "nexus of contracts" has significant implications:[24]

1 The firm *per se* has no veritable existence, and it makes no sense to speak of such things as the behaviour or interests of a firm. Furthermore, and this point is crucial, the question of who owns the firm doesn't even arise, as Fama clearly upheld:

> Ownership of capital should not be confused with ownership of the firm. Each factor in a firm is owned by somebody. The firm is just the set of contracts covering the way inputs are joined to create outputs and the way receipts from outputs are shared among inputs. In this "nexus of contracts" perspective, ownership of the firm is an irrelevant concept.
>
> (Fama 1980: 290)

Likewise, of course, we can never consider the existence of contracts entered into *by the firm itself* (with employees, managers, other firms, etc.), or the idea that it is the firm which owns assets.

2 The question of the frontiers of the firm is meaningless:

> It makes little or no sense to try to distinguish those things that are "inside" the firm from those things that are "outside" of it.
>
> (Jensen and Meckling 1976)

In particular, this means that the employees are not to be considered as part of the firm, any more than the suppliers of raw materials, or the customers.

3 And, quite logically, there is no fundamental difference between the firm and the market – except perhaps in the characteristics of contracts. This point was expressed very clearly by Alchian and Demsetz (1972) when, unlike

> Simon (1951), they rejected the idea that contractual relationships specific to the firm can imply any sort of "authority" relationship different to that which can be found in market relationships, a point of view subsequently adopted by Jensen and Meckling (1976).[25]

One may reasonably wonder whether, ultimately, these conclusions do not apply to all strictly individualist and contractual conceptions of the firm, which always tend to reduce the analysis to bilateral relations between people – with, however, one key question: that of the characterisation of the employment relationship and the employment contract. This aspect involves the contents of the relationship between the managers – or more precisely, in fact, the firm itself – and the workers; it is indissociable, in our opinion, from the issues concerning the internal organisation of production. No doubt, we can best apprehend the "nature" of the firm by combining recognition of the existence of an authority relationship, expressed *particularly* in the specificities of the employment contract, with the specific character of the organisation and implementation of production within the capitalist firm – which lies outside the scope of the contractual approach.

Let us now return to the assertion that the manager is the exclusive agent of the shareholders. Quite apart from the arbitrary nature of this principle – one could certainly argue that it actually represents no more than an initial convention having the character of a *constitutive rule* –it is not unreasonable to suggest that there may be a fundamental contradiction in the reasoning presented above. Reducing the firm to a nexus of contracts, especially with the suppliers of the different factors of production, combined with the assertion that no one can be considered the "owner of the firm," implies that the shareholders are the owners *of the capital*,[26] that is to say, of one of the factors of production, but that they cannot, under any circumstances, be considered the owners *of the firm*. In this case, there is no reason to consider the managers, whose function is to ensure the coordination of production and of all the firm's activities, including the relations with all the different suppliers of factors of production and with other contracting parties such as buyers, as being the exclusive agents of the shareholders. They are in fact the agents of all the stakeholders. The most coherent way to describe the situation would be to say that we are in a principal–agent relationship, but with *several principals*. This leads us to consider the firm essentially as a coalition, which is, to a degree, the viewpoint expressed in managerial theories, or in the "stakeholder" conception of corporate governance. And this echoes *certain aspects* of Berle's and Means's viewpoint.

In the vision of the firm as a nexus of contracts, it is hard to see what can justify, *a priori*, the special status granted to shareholders, unless the reasoning is based on something other than ownership. The most common argument is based on the question of risk-sharing: the shareholders are paid on the basis of "residual income" (they are "residual claimants"), and are hence expected to bear the risks. This interpretation is extremely fragile: it needs to be supported by rigorous analysis of the character and operating conditions of the different

markets for "factors of production," the conditions governing the sharing out of the surplus created by the firm's activity and the modes of payment of the different parties involved, in other words the exact design of the different contracts. It would be rash, to say the least, to suppose that the shareholders bear all the risks, given the present operating conditions of labour and capital markets. We must not forget that the liquidity of financial markets fundamentally changes the position of capital providers. And it would be more realistic to consider that the "residual income" is not that of the shareholders, but that *of the firm itself*. But this would mean recognising, to the contrary of agency theory, that the firm is a specific entity with a distinct income, which entails a totally different conception of the corporation.[27]

An alternative view does indeed exist, within the perspective of the "stakeholder" conception of corporate governance. Here, the emphasis is laid on the organisation of cooperation between the different members of the firm, the function of the managers being to manage the relations and potential conflicts of interest between the stakeholders and to ensure the durability and expansion of the firm. The "partnership model" explored by Radner (1986) can be situated within this perspective. Zingales' reflections on the transformation of the capitalist firm, which led him to maintain that the new firm is more like a cooperative,[28] move in a similar direction. Blair and Stout (1999) propose a theory of the public corporation, an alternative to agency theory, in which they explore this conception in great detail. Their reasoning sheds much light on the problems raised by such a view. Their point of departure lies in the emphasis placed on the problems of *horizontal* coordination between the different members of the "team" that is constituted by the firm: the central function of the hierarchy is to ensure this coordination and to manage the potential conflicts between team members, "to mediate horizontal disputes among team members," (Blair and Stout 1999: 264). This implies that the firm is viewed from the start as a constituted team, and that the members of the team are precisely defined. For Blair and Stout, these are the shareholders, the employers and the managers, the latter occupying a specific position, as is the case with the "monitor" of Alchian and Demsetz. The firm then appears as a "coalition," which, in a certain manner, maintains a contractual perspective, as we shall see.

This emphasis on the firm as "production team" leads to a new dimension being taken into consideration, a dimension neglected in the analyses we have examined so far, which have been focused on the vertical relations between employers and workers, or between shareholders and managers. This new dimension involves the *horizontal* relations and interactions between the different agents participating in the activity of the firm (the team members) – interactions on which the efficiency of production and the level of surplus created are dependent. We are, in other words, at last dealing with a genuine investigation into coordination within the firm. This leads to consideration of the collective dimension of the firm's activity, the fact that "the essence of team production is that the whole can be made bigger than the sum of parts" (Blair and Stout 1999: 269). This is what Blair and Stout do, but as they remain within a contractual

perspective and within a standard framework of rational choice, they tackle this question in a particular manner: the problem of coordination is reduced to the problem of managing free rider behaviour. This problem is considered all the more significant as, following Rajan and Zingales, the authors stress the importance of "firm specific" investments made by the team members, creating irreversible interdependencies between them.[29] As with Alchian and Demsetz, the solution to this problem is supposed to reside in the recourse to a third party. This leads to the idea that it is all the members of the team – including the shareholders and employees – who decide, by common assent, to confer the control rights over the firm's production and the use of inputs to an "outsider," namely the hierarchy. In this, the authors explicitly express a "Hobbesian" conception of the firm which, as Bowles (1985) observed, lies at the heart of contractual approaches. The public corporation with dispersed share ownership – the analysis clearly applies to this sole case, and not to the case of an unincorporated or single-shareholder company – is thus not so much a set of material assets under unique control (as in property rights theory) as an agreement between individuals aiming to work together for the benefit of everyone. The mode of organisation is chosen essentially with a view to guard against opportunism and free rider behaviour.

The authors set out to show that, in the United States, the law tends to confirm this view, rather than the view proposed by agency theory, particularly in so far as directors are "trustees more than agents" and the corporation is actually considered a "legal person," and therefore an entity distinct from its founders or shareholders. They seek to show – though one may ask whether this is the necessary consequence – that the directors are supposed to act in the interests of this entity, and not solely in the interests of the shareholders.

This is not the place to launch into a discussion on the contents and implications of American corporate law. Certainly, the major changes in law should be taken into account, but so should the changes in all the formal and informal rules and norms governing the functioning of companies, and more particularly public corporations over the last twenty years, in the wake of debates on corporate governance and the trend towards the "return of the shareholder." Whatever the case, the fact remains that the representation proposed here might inspire a democratic or cooperative conception of the firm, but that given the essential characteristics of the structure of the large corporation – especially from a legal point of view – and the structure of our economic systems, it remains essentially a "contractual fiction." This view should result in managers being accountable, at all levels, to the team members who are supposed to have chosen them.[30] This would require a contractual (and legal) system, and a system of rights, totally different to those which prevail in corporations. If we return to the effective characteristics of the hierarchical system in the capitalist firm: at every level, those in charge are accountable to the level above, and, of course, the main constituent feature of the public corporation is the fact that the managers are supposed to account first and foremost to the representatives of the *shareholders*.[31] It remains to be explained how this can be justified

economically and considered compatible with the conception of managers as being an expression of the will of all the team members, including the employees.

Finally, we must point out that, although Blair and Stout, like Rajan and Zingales, attach central importance to the analysis of actual production and the conditions under which the firm produces a surplus, they do not give themselves the means to really tackle this issue, never venturing beyond the framework of the questions raised by the contractual approach: how to construct contracts, and if necessary allocate ownership, so as to obtain an efficient form of contractual equilibrium in the face of egoistical and opportunistic behaviour. Indeed, their analysis of the team, like that of Alchian and Demsetz, only takes into account the relationship between the team members on the one hand and the manager-monitor on the other; it neglects the relations and interactions between the team members. In short, it says nothing about the organisation of work and production. The analysis of production, of the organisation and development of the division of labour and of cooperation within the firm, remains the missing dimension. Yet it is far from being the least important.

## Production and capabilities: organisational architecture and the knowledge-based and evolutionary perspectives

The problems of the organisation of production, in connection with the implementation of individual and collective skills, had already been evoked in the old reflections on capitalism, particularly in Smith and in Marx's exploration of manufacturing and large industry, drawing on the works of Ure or Babbage. Broadly speaking, the problems of the organisation of production can be seen as consisting in the simultaneous conception of a mode of division of labour and of coordination between the tasks thus defined. More specifically, attention can be focused on the problem of coordination caused by the dispersion of information and knowledge between individuals, an aspect which, as we have seen, is one of the basic elements of the agency theory of Jensen and Meckling. The "economic theory of teams" developed by Marshak and Radner (1972) proposed a first exploration of this type of question. This work is important, in that it introduces a representation of the firm and of organisation, especially hierarchy, radically different from that presented in contractual approaches.[32]

A "team," as defined by Marshak and Radner, is a collective composed of individuals who share *the same objective*. By adopting this hypothesis, we set aside the problems resulting from conflicts of interest between individuals and groups, to concentrate on the problems of coordination arising from the fact that individuals do not possess the same information. The firm, like any other organisation, is built on a *decentralisation of information*, made necessary by the bounded rationality of individuals and their limited information-processing capacities, and thus on what we can describe as a *cognitive* division of labour. So, one of the main features of the modern firm is that it is built on a system of information, *distinct from the price mechanism* (Aoki 2001: 95).[33] The objective of team theory is to determine the optimum mode for the decentralisation of information (Radner

1992) and to explore the efficiency properties of different forms of organisational architecture. Efficiency depends on a trade-off between "the number of processors," i.e. the number of decision centres on the one hand, and "the delay between the receipt of information by the organisation and the implementation of decisions" on the other (Radner 1992: 1400). One of the results of this approach is to demonstrate that efficiency, so defined, can be achieved by an organisation based on "hierarchical networks." This endows the concept of hierarchy with a meaning and justification significantly different to those we have considered up until now. Here, hierarchy describes a mode of organisational architecture, with an arborescent structure, built around asymmetrical vertical relationships between sub-systems and decision centres, the purpose of which is to optimise the management of information. This architecture corresponds to the typical structure of the management systems of large companies. Hierarchy thus involves the way in which the division of labour – and more particularly cognitive labour – and decision-making are organised within the firm. Within this approach, one can identify different types of organisational architecture and possibly consider the specific problems of governance they create, as Aoki has done (2001: chs 4 and 11). The evolutionary approach, which we shall examine later, tackles the same types of question, by investigating the conditions of decomposability of a problem into sub-problems, following the perspective opened by Simon (1996), and by taking learning processes into consideration (see especially Marengo and Dosi 2005; Marengo *et al.* 1999).

Before going any further, we should note that this type of analysis of the organisational architecture of the firm needs to be completed by investigating the conditions under which a given architecture is *viable*. This involves taking individual and collective behaviour into account and, more broadly, recognising that "organizational architecture needs to become institutionalized," as Aoki puts it (Aoki 2001: 97). What does this entail? We shall return to this point later.

What we are offered on this basis is a picture, in many respects more realistic than the contractual view, of what the organisation of a large company actually is. This organisation is marked, as Chandler points out and as Radner (1992) recalls, not only by the existence of a manager and an authority relationship, but also by a complex apparatus of management, which constitutes a key component of the firm, the expression of a certain mode of division of labour (founded, let us note in passing, on the separation between decision and execution). This apparatus of management constitutes precisely the "visible hand" of managers, of which the characteristics, the different modes of structuring and the genesis were brought to light so brilliantly in the works of Chandler: systems of administrative coordination based on a hierarchical system and centralised control. This organisational architecture ensures the economic unity of the firm; it is through the formation of this system that the modern firm confirms its existence as an economic entity. Contractual and property relationships do exist, and should not be overlooked, but they only represent one among several supports of this system. No less essential, the accounting system constitutes a key element in the system of information and management, and thus contributes to

the constitution of the firm's unity, as does the whole set of formal and informal rules, procedures and routines around which the system is organised and reproduced.

If we now return to the foundations of this representation, one essential aspect must be taken into account: it is the *bounded* rationality of individuals which lies – at least partly – at the origin of the need for the cognitive division of labour and the dispersion of information and knowledge. This demands a profound rethinking of the analysis of behaviour, and consequently of the approach to organisations, in relation to what the economic theories of rational behaviour underpinning the contractual approaches have to offer. And this is exactly the point on which "competence-based" and evolutionary theories of the firm are founded.

The literature adopting this perspective is abundant and varied, both in its origin and in the nature of the questions it tackles. Its origin lies in works in the field of strategic management, defining a so-called "resource-based" perspective,[34] pioneered by Penrose (1959) and the works of Chandler. These works converge with the evolutionary theorisation of the firm and organisations, developed from the work of Nelson and Winter (1982). It is here that we can find the most systematic attempt to build an alternative theory to the contractual vision, aiming to explain the structure and behaviour of organisations as the consequences of individual and collective learning dynamics.[35] The question these theories initially seek to answer is the following: "why do firms *persistently* differ in their characteristics, behavior and performances?" (Dosi and Marengo 1994). Broadly speaking, the answer is that firms differ in the nature of the specific knowledge and competencies they are capable of producing and accumulating, and which represent their critical resources. Thus, the unifying characteristic of these analyses is that they consider the firm fundamentally as a "bundle of capabilities" rather than a "nexus of contracts." They have three main features in common:

1 They consider the firm first as a place of *production*. In particular, this means that the elementary unit of analysis is the task, not the transaction (or contract), one question then being to explain how the division and coordination of tasks in the firm is achieved (Marengo and Dosi 2005). This steers the analysis of production in a particular direction, essentially *cognitivist*: the activity of production is first and foremost a question of "problem-solving." One consequence of this approach is that technology and the division of labour are no longer considered as givens – as they are in all the contractual approaches, including Williamson's transaction costs theory – but on the contrary as the results of organisational choices and choices in modes of governance, or at least as elements that evolve in interaction with these choices.
2 This leads to the second major feature: the central position given to the question of knowledge and learning, in contrast to the emphasis laid on information in the contractualist viewpoint. This analysis is based on an

economics of *knowledge*, radically different to the economics of *information*.[36] The shift from a theory founded on information to a theory founded on the concept of knowledge represents a break that has two major dimensions. The first of these involves the importance given to *tacit* knowledge and to knowledge *specific* to the organisation. This point is well-known; it is central to every analysis of organisational capabilities and learning. The second dimension is even more important: knowledge can be distinguished from information by the fact that it doesn't have the status of a simple description of a given reality, it constitutes a *constructed representation*. Knowledge is a social construct that includes a dimension of belief and judgement; knowledge is a "justified true belief " (Nonaka 1994). This aspect is central to the understanding of learning processes and, above all, collective learning. It can be found in all the variants of organisational learning and evolutionary theories: at the heart of learning problems reside the conditions under which the representations, languages and framing processes are formed, the effects of these representations and the difficulties in changing them.

Here, we arrive at the very heart of the competence-based view of the firm. This means that the whole analysis is based on behavioural and micro-economic foundations that represent a radical departure from traditional micro-economics, even in its most recent developments (information economics and game theory). This approach gives a central role to the analysis of rules, routines, conventions and shared representations as elements that structure organisations and institutions, and the firm in particular. It also leads on to the third main feature:

3 These analyses adopt a dynamic perspective that breaks with the search for static equilibriums. This dynamic perspective shapes both the analysis of behaviour and interactions between agents and the study of organisations and institutions. The problem no longer lies in determining which, out of several different organisational forms and modes of governance, will be the most efficient in a given context: now the objective is rather to understand the conditions of emergence and transformation of these forms and the properties of their dynamics of evolution, likely to be marked by strong path-dependency.

We cannot discuss these theories in detail here; we shall simply attempt to describe the contribution they make to our understanding of the firm, as alternatives or complements to the prevailing contractual viewpoint.

How can we characterise the competence-based, evolutionary firm? The answer is not straightforward: "Business organizations are *behavioral entities*, characterized by seemingly quite different arrangements in terms of the cognitive and operational division of labour, as well as equally different hierarchical set-ups, patterns of access to information and incentive and control structures" (Dosi et al. 2003: 2). The firm is considered as (i) a specific entity, whose behaviour can be observed and analysed, and (ii) an entity characterised by two structuring principles: the first involving the organisation and division of

labour and the activity of actual production, and the second involving all the systems and procedures of incentive and control. Whence two categories of questions, concerning the problem-solving features of organisations on the one hand, and the problems of confrontation between the divergent aims and interests of the individuals and groups which constitute the organisation on the other. Coordination through organisation (and coordination in general) is considered to have two dimensions: first, "*cognitive* coordination," focused on the constitution of a collective ability to solve problems – and so to produce and innovate – and second a "*political* coordination," the aim of which is to resolve, or at least to regulate, conflicts of interest.[37]

To all intents and purposes, the contractual approaches adopting a "*pure incentive-governance view*" (Dosi *et al.* 2003: 5) only tackle the second of these dimensions, *and they do so in their own fashion*, i.e. on the basis of hypotheses of the economic theory of rational behaviour and the search for Pareto optima, while evolutionary and competence-based analyses concentrate on the first dimension,[38] and do so, as we have already seen, on the basis of radically different behavioural hypotheses. Thus, the competence-based and contractual visions of the firm are opposed in terms of both the questions they tackle and their theoretical foundations. The trouble is, these two categories of questions are closely interlinked. We believe that one of the central problems in the theory of the firm lies in understanding how these two dimensions are connected. We shall return to this point later.

In the meantime, what should hold our attention here is the idea that the firm represents a coherent system of *organisational capabilities*. These capabilities are essentially constructed through a long internal process of collective learning, marked by strong path-dependency. The heart of the evolutionary and competence-based analysis of the firm is thus centred on categories of learning, organisational capabilities and routines, these routines constituting – for the evolutionists – the central instrument of cognitive coordination within the organisation (see Coriat and Weinstein 1995, 1998).[39]

Here, we must consider certain aspects of these organisational capabilities that are emphasised by the authors, and that are important for understanding how, in this approach, the firm forms a coherent whole: *skills and organisational learning are fundamentally social and collective.*[40] This collective dimension has several characteristic features (Weinstein and Azoulay 1999):

- The knowledge and capacities of the organisation are greater than the sum of the knowledge of its members. *The greater part of its capabilities is specific to the organisation*: "it is the firms, not people that work for the firms, that know how to make gasoline, automobiles and computers" (Winter 1982).[41] The skills of the firm exist, replicate and develop, even though the individuals that make up the firm change and its structures evolve.
- The knowledge and skills of each member of the organisation are fragments of the knowledge of the organisation, fragments which can only assume their significance and effectiveness within a certain context, constituted by the

whole set of knowledge and skills of the other members and the relations between them (Winter 1982). Individual knowledge and skills are thus "firm-specific."

- The processes of learning and the production of knowledge, individual or collective, are social processes, incorporating the dimensions of training and the use of "cognitive frames," of *shared* codes, beliefs and representations. This raises the question of the conditions of coordination of individual knowledge and skills, and the conditions that make collective learning possible: the constitution of a common knowledge base, of shared representations and languages, of performance assessment systems, of coordinated problem-solving procedures. In particular, any form of durable organisational learning requires mechanisms for the *codification of knowledge and interaction procedures* (Dosi et al. 2003).
- Thus, the structure of the organisation and the system of relations of which it is composed are essential components, and a necessary condition of organisational capabilities. The firm's capabilities are embedded in its social structure: its organisation, rules and routines, culture, and so on (see Hodgson 1998). In this sense, it is the whole constituted by the organisational architecture, the interaction procedures, the rules, routines and representations, and the collective capabilities that are produced by them, which define the unity and identity of the firm as a productive unit.[42]

The contribution of the competence-based approaches is essential for grasping the essence of the firm, its unity and identity. We believe this to be even more true when the analysis of capabilities is combined with the analysis of organisational architectures and management systems, which can be found in Chandler's work.[43] Two big questions remain, involving relations with contractual issues and with the subject of institutions.

As we have already remarked, analysis of the firm is structured around two questions, concerning what we have called cognitive coordination and political coordination. In our opinion, one of the main problems with competence-based analyses is that they gloss over the second question, by implicitly assuming that the formation of collective capabilities and collective learning can be explained without reference either to the modes of resolving conflicts of interest or to the systems of incentive and control.[44] Yet the two issues are obviously closely interlinked: the forms of the division of labour and the distribution of competencies in the firm are central to the conditions of organisational learning. At the same time, they define positions, "roles" and structures of power and interest. The modes of managing the work force, within the firm and in the labour market, play an essential role in the mobilisation of individual skills and the formation of organisational capabilities, while also, of course, forming a fundamental dimension in the system of incentive and control. In the same manner, as we have seen, the hierarchy can be conceived both as a mode of division of labour intended to optimize the management of knowledge, and as a mode of incentive and control. Likewise, taking into account the "political" dimension of the firm

requires us to rethink the nature of organisational routines and the analysis of skills; this is what Dosi and Coriat demonstrate clearly in their analysis of Taylorism: "routines not only represent problem-solving procedures, but are at the same time control and governance devices."[45] Taking into account the double dimension of routines, and more generally the whole set of organisational procedures and forms, leads to a profound transformation in the analysis, compared to a purely cognitive approach. And it is because of their single-minded focus on the cognitive dimension that evolutionary and competence-based theories tend to have little or nothing to say about the fundamental relationships around which the capitalist firm and the corporation are structured: the employment relationship, the relations between shareholders and managers and the role of property relations.

The incorporation of contractual dimensions and forms of incentive and control on the one hand, and of productive and cognitive dimensions on the other, is certainly one of the biggest challenges facing the theory of the firm. But it can only be surmounted by adopting unified behavioural hypotheses, in other words by rethinking questions of governance, incentive and control, particularly by drawing on recent advances in cognitive theory, following on from the theses of Simon and of behavioural economics.

A second major question involves the specifically institutional dimension. Competence-based analyses generally remain profoundly "organisational," in the sense that they focus on the *internal* interactions between members of the firm, which is portrayed as a "closed" system, just as contractual approaches portray it as a "private" system. The emphasis placed on the accumulation of knowledge through *internal* learning, to the detriment of the acquisition of external capabilities and knowledge, leads in this direction. If we go no further than that, we risk overlooking the role played by the social and institutional environment in determining organisational forms and constructing the unity of the firm. These forms of organisation, like the routines and modes of learning and knowledge accumulation, are conditioned by institutional forms – especially national ones – such as those which regulate labour markets, capital markets, intellectual property or accounting systems. The examples of "Taylorism" or "Ohnism" also show the importance of the social processes through which modes of organisation, including systems of routines, can be codified and disseminated on a national and international level. But there is more: to take the *political* dimension of the firm fully into consideration we must move beyond the economic vision focused on a pure logic of static or dynamic efficiency, characteristic of both contractual and cognitivist approaches, to grasp the importance of the dimension of *legitimacy* inherent in the rules which found the firm's structure and operation. Only the specifically institutionalist or "conventionalist" approaches can achieve this.[46]

More fundamentally, taking into account the institutional dimension of the capitalist firm leads us to recognise that the firm is not just a simple private agreement, it is also a component of a certain social system, and this is the only way fully to grasp its specificities and nature.

## By way of a conclusion: beyond contracts and competencies, understanding the firm as an institution

The view offered by the economic theory of the firm remains fragmented: there is a separation between the Coasian approach and that derived from Berle and Means; there is a separation between the contractual approaches and those which take production and skills into consideration, to give just a few examples. And so the identity of the firm, when it is recognised, appears to be multiform. As Chandler observed (1992: 79), the theoretical literature highlights at least four attributes: the firm is a *legal entity*, an *administrative* (or managerial) *entity*, "for teams of managers must coordinate and monitor its different activities," what we can describe as a *productive entity*: "a pool of physical facilities, of learned skills and liquid capital," and lastly "the primary instrument in capitalist economies for the production and distribution of current goods and services and for the planning and allocation for future production and distribution." To these four, we can add that the firms exists, in our current system, as a *financial entity*, particularly as a bundle of (non-human) assets, subject to a unified control right, reflecting the fact that if the capitalist firm is an instrument for the production of goods and services, it is also – and primarily – an instrument for the valorisation of capital. Finally, the firm is the central actor in the dynamics of capitalism, as Schumpeter and Perroux have observed.[47]

These different viewpoints may express different theorisations, but they also reflect the very nature of the capitalist firm. The firm as entity – and as the central institution of capitalism – is the union of all these different attributes,[48] but this union is in permanent re-composition: the relations between its legal foundations and legal identity, its management and production system and its organisational architecture, its financial profile and structure and its modes of dynamics all evolve with its historical transformations and with the construction of specific national (or multinational) and sectoral forms.

This union can be understood by considering the capitalist firm as an institutional form, a component of an institutional system. This means first that the firm cannot be reduced to a "private agreement," i.e. the simple result of an agreement between individuals, either one-off or long-term, with freely-defined contents. It is, first and foremost, a public entity, a component of a certain social order. The explicit or implicit contractual relations between individuals in the firm – which are in a continual process of formation and dissolution – are built on a certain number of *constitutive rules* that found the capitalist company, as Commons, among others, has shown.[49] These rules define the fundamental relations that constitute the firm. Thus, the firm is "instituted": it is built on "external" rules (external, that is to say, to the specific rules, norms and routines constructed for itself by each firm). These external rules are imposed on individuals. They found the legal existence and identity of the firm, they define the fundamental characteristics governing its internal and external relations and modes of operation, and they condition its structures of power and governance. It is at this level, first of all, that the firm is established as a specific entity.

These constitutive rules are those of a certain institutional order, a set of formal and informal social rules, imposed on everyone, the product of collective actions of a fundamentally political nature. They govern and constrain individual actions and interactions – and therefore private contractual forms – while at the same time permitting the "liberation," to use Commons's expression, of individual actions and allowing the formation of private agreements. Within this institutional order, in our societies, the legal system occupies a central position, as Commons showed and as Coase recalled, particularly by defining the legal persons who can carry out transactions and enter into contracts (Hodgson 1999: 236). In this respect, the legal dimension is constituent of the firm, if only, in the first place, because of the recognition of the firm, then the public corporation, and ultimately even the financial group, as legal entities. In this legal system, the definition of property and its forms itself occupies a central position. Commons (1924–95) showed how the emergence of the capitalist firm depended on transformations in private property, "by means of the enlargement of the concept of property from things to the exchange value of things" (Commons 1924–95: 157), which appeared with intangible property and the concept of "assets,"[50] while Berle and Means emphasised the *de facto* transformation of the contents of property and its "depersonalisation" by the development of capital markets and transformations in the public corporation.

The accounting system and its basic principles are also a main component, and the expression of the institutional order of the capitalist firm. This system is built up through the action of the firms themselves, but is then "instituted" through the spread of accounting practices and conventions and their establishment as legal constraints.

This institutional system which founds the firm has two main dimensions: one involving the employment relationship, and one involving corporate governance. These relationships and the principles that govern them are more than contractual rules and systems; they are social relations that reflect the structure of society and the institutional complementarities and hierarchies around which society is organised. Thus the employment relationship is built on the property regime, on the way in which the large company is structured (especially the modes of division of labour it adopts) and on the formal and informal rules which govern the constitution and operation of labour markets on the one hand and the conditions of the use of labour within the firm on the other. Likewise, the principles of corporate governance and the status of managers are based on the development of capital markets and the rules and institutions that govern them, on the modes of relationship built up between banks and industries (particularly in countries like Germany and Japan), on changes in the very definition of property and its forms and, once again, on the internal transformations of large companies.

There are certain constituent rules or "conventions" that come into play in the construction of these institutional orders. They serve as a foundation, and above all a justification, for certain aspects of the legal or informal norms that regulate the corporation, among others. These conventions are basic principles

governing the mode of governance and the status of the different stakeholders. This is the case, for example, for the affirmation of the principle that the managers are the agents of the shareholders, the latter being considered as owners of the capital or, on the contrary, the principle that the managers are the expression of the will of all the stakeholders, and the trustees of a certain collective interest. It is also the case for the basic principles influencing the definition of accounting rules (especially as regards the mode of evaluation of the different types of assets), which express a certain conception of what the firm is. It could also be the case for the extent to which the employees are considered part of the firm. The status of the workforce under the Fordist regime was largely based on its recognition as a constituent of the firm. This recognition grew out of transformations in the firm itself, marked by much stronger integration of the workers – related to the construction of hierarchical management systems, and thus with the affirmation of the firm as a specific entity – and out of a gradual change in the status of the workers. This change, notably expressed in labour law, contributed to the construction of a new constituent convention, certain aspects of which were obtained through industrial and political struggle.

Whence the importance of the social and political conditions governing the construction, transformation and imposition of these fundamental principles that are supposed to govern large social institutions in general and the firm in particular. Today, this importance is expressed in all the debates and conflicts over such issues as the definition of the principles of corporate governance, changes in labour law or intellectual property. And in is in this way that the "nature" of the firm is socially constructed and transformed.

## Notes

1 See also Simon (1991, reprinted in this volume).
2 One group stands out by its absence: the consumers, or more generally the users. Modern theory of the firm is in some ways the opposite of the conception of the firm in traditional neo-classical theory, which focuses on the relation between production and *demand*.
3 On this point, see Kirat (this volume) and Berle (1947, Reprinted in this volume).
4 On what differentiates transaction costs theory and property rights theory, see Gibbons (2005), and, emphasising the contrasts, Williamson (1991, 2000) and Kreps (1996).
5 On this point, see Tirole (1999).
6 See Grossman and Hart (1986); Hart (1989), who identifies the firm with "all the non-human assets that belong to it, assets that the firm's owner possesses by virtue of being the owner of the firm," or Hart and Moore (1990): "we identify a firm with the assets that its owners control," and, later: "we identify a firm with the assets it possesses."
7 Even though it appears that the individual is assumed to be the owner of the assets, to the extent that he is the owner of the firm, in which case we need to know the foundations of this ownership, and what it means to be the owner of a firm. We suggest that it is up to the law to settle such questions.
8 See Kirat (this volume) and Manfrin (this volume).

9 It should also be noted that effective power, so defined, will depend on the state of the labour market, a point seen clearly by the American radicals, in their analysis of "contested exchange" (Bowles and Gintis 1993).
10 On this point, see Hodgson (1999: ch. 10).
11 See Kirat (this volume).
12 On which Williamson draws explicitly. On this point, see Gibbons (2005).
13 Undertaken notably by Commons (1924; 1932).
14 Because of transformations in the "nature of the firm" connected with more general transformations in capitalism (Rajan and Zingales 2000).
15 Or again: "in our framework, the firm . . . is a collection of commonly owned critical resources, talents, and ideas, and also the people who have access to those resources" (Rajan and Zingales 1998: 404–5).
16 The resource-based theory of the firm "does not indicate how these resources, especially intangible ones, provide authority around which organizations can be built" (Rajan and Zingales 2001b: 842).
17 From a different perspective to that of Berle and Means. But the two approaches are complementary, through the questions they raise about the role of property in the structuring of the firm and of the wider economic system.
18 Which brings to mind the thesis of the radical Americans that the capitalist firm is founded on a system of inefficient property rights (Bowles and Gintis 1993).
19 Berle and Means (1932: 309) quote a particularly significant passage by Walther Rathenau: "No one is a permanent owner. . . . This condition of things signifies that ownership has been depersonalized. . . . The depersonalization of ownership simultaneously implies the objectification of the thing owned. . . . The depersonalization of ownership, the objectivation of enterprise, the detachment of property from the possessor, leads to a point where the enterprise becomes transformed into an institution which resembles the state in character."
20 Berle (1947) returned to this point, showing how the firm as an entity continues to exist, and to be legally recognised, beyond the legal identity of each company, when networks of financial relations and larger groups develop.
21 See especially Jensen and Meckling (1992–98). The theses of Hayek (1945) are used here, particularly the distinction between specific and general knowledge. It is reasonable to wonder whether the Hayekian view doesn't entail a profound rethinking of the standard principles of rational choice which remain the foundation of agency models, and of the majority of contractual analyses and models of incentive. We shall return to this point.
22 What Jensen and Meckling, and property theorists, have to say should be considered in the light of historical transformations in the concept of ownership, highlighted in particular by Commons (1924) and Berle (1965, reprinted in this volume).
23 A theory which preceded that of Grossman, Hart and Moore and which has much wider objectives as to the analysis of property. See, for example: Alchian (1987), Demsetz (1967, 1988), Barzel (1997).
24 On this, see Coriat and Weinstein (1995: ch. 3).
25 In the same spirit, see, Cheung (1983).
26 The concept of capital implicit here must be precisely defined: the shareholders are suppliers of funds; to remain within a coherent contractual perspective, they simply have a different type of contract from that of suppliers of capital paid at a fixed rate. They are not owners of the capital "of the firm," a term which has no meaning in this context.
27 Which can be found, for example, in the work of Perroux (see Kirat, this volume), Zappa (see Canziani, this volume), or, more recently, Blair and Stout (1999). Biondi (this volume) further develops the links with accounting, from this perspective.
28 See Zingales (2000: 1648): "traditionally one of the great advantages of the corporate structure was that it concentrated all the control rights to a group with very homogeneous interest: the shareholders. . . . The recent changes in the nature of

the firm . . . make new firms more similar to cooperatives, rather than to traditional corporations."

29 The public corporation is "not so much a 'nexus of contracts' (explicit or implicit), than a 'nexus of firm-specific investments', in which several different groups contribute unique and essential resources to the corporate enterprise, and who each find difficult to protect their contributions by explicit contacts" (Blair and Stout 1999: 275).

30 "*Hierarchs work for team members (including employees)* who 'hire' them" (Blair and Stout 1999: 280, italics in original).

31 Even if managers are also subject to a certain number of formal and informal institutional constraints which govern their actions, embodied, for example, in regulatory bodies such as the SEC (in the United States), anti-trust authorities or consumer pressure groups. On this point, see Berle (1965, reprinted in this volume).

32 Although it is possible to combine the two approaches, as Radner (1986) and Aoki (2001) have done.

33 Which includes, in particular, the accounting system. See Biondi (this volume) who develops a similar idea on the basis of the accounting system.

34 See, amongst others, Wernerfelt (1984), Rumelt (1984), Montgomery (1995), Teece *et al.* (1997). See Foss (1997) for a re-issue of major contributions.

35 See especially Dosi and Marengo (1994), Dosi et al. (2000), Dosi, Levinthal and Marengo (2001), Dosi, Faillo and Marengo (2003), Marengo and Dosi (2005). Other works, drawing on Richardson in particular, also seek to reconstruct a theory of the firm from a similar perspective (see Langlois and Robertson 1995).

36 On this point, see for example Cohendet (1998); Weinstein and Azoulay (1999).

37 This political dimension is central to authors such as Berle and Means, or Commons.

38 Certain works do attempt to combine the two categories of questions, but they remain scarce and undeveloped. See: Coriat and Dosi (1998), Dosi *et al.* (2001).

39 The concept of routine, and the relations between routines and skills, lie at the centre of the evolutionary view. Here, we set aside this question, which deserves detailed investigation. See Cohen *et al.* (1996); Dosi *et al.* (2000).

40 "Learning processes are intrinsically social and collective and occur not only through the imitation and emulation of individuals . . . but also because of joint contribution to the understanding of complex problems" (Teece and Pisano 1994: 544).

41 Which does not prevent Winter from espousing a strict methodological individualism.

42 This appears clearly in the definition of the firm's capabilities given by Chandler (1990: 594): "These organizational capabilities were the collective physical facilities and human skills as organized within the enterprise. These included the physical facilities in each of the many operating units – the factories, offices, laboratories – and the skills of the employees working in such units," or even more clearly in the description by Leonard (1998: 19), who identifies four dimensions to the firm's capabilities: "Employee knowledge and skills"; "Physical technical systems"; "Managerial systems"; "Values and norms."

43 Chandler (1990, 1992) explicitly champions the competence-based theory of the firm against the transaction costs theory. However, it is questionable whether his *historical* vision is really the same as the *cognitive* vision of the evolutionists. On this point, see Coriat and Weinstein (1995: ch. 4).

44 This is what Nelson and Winter do, with their "truce" hypothesis. A small number of works seek to go further (Coriat and Dosi 1998; Dosi *et al.* (2001).

45 See also Coriat and Weinstein (1998): "we do not have 'cognitive' rules and routines on the one hand and 'disciplinary' rules and routines on the other . . . , it is the same routines or systems of routines and the same organisational forms that structure the dynamics of learning and the processes of conflict management and allocation of the surplus."

46 Which we can find in the works of Berle and Means, Perroux (see Kirat, this volume), or Commons (see Bazzoli and Dutraive 2002). See also Eymard-Duvernay (2004).

47 See Kirat (this volume).

48 See Biondi (2005), who presents another way of analysing the firm as entity, as the interconnection of different dimensions.
49 See Bazzoli and Dutraive (2002).
50 Which "includes every thing that can be sold" (Commons 1924: 158): material goods, plant, land and buildings, but also all sorts of financial assets, patents, copyright and trade-marks, "and even the goodwill of the business."

## References

Aghion, P. and Tirole, J. (1997), "Formal and Real Authority in Organizations," *Journal of Political Economy*, 105(1): 1–29.

Aglietta, M. and Reberioux, A. (2004), *Dérives du capitalisme financier*, Paris: Albin Michel. English translation (2005): *Corporate Governance Adrift. A Critique of Shareholder Value*, Cheltenham: Edward Elgar Publishing.

Alchian A.A. (1987), "Property Rights," in J. Eatwell, M. Milgrate and P. Newman (eds), *The new Palgrave: A Dictionary in Economics*. London: Macmillan.

Alchian, A.A. and Demsetz H. (1972), "Production, Information Costs, and Economic Organization," *American Economic Review*, 62: 777–95.

Aoki, M. (2001), *Toward acomparative institutional analysis*, Cambridge (Mass.): MIT Press.

Baker, G.P., Gibbons, R. and Murphy, K.J. (2002), "Relational Contracts and the Theory of the Firm," *Quarterly Journal of Economics*, 117(1): 39–84.

Barzel, Y. (1997), *Economic Analysis of Property Rights*. Second Edition. Cambridge: Cambridge University Press.

Bazzoli, L. and Dutraive, V. (2002), "*L'entreprise comme organization et comme institution: Un regard à partir de l'institutionnalisme de J. R. Commons*" [The enterprise as organisation and as institution: a view from the institutionalism of J.R. Commons], *Economie et Institutions*, 1: 5–46.

Berle, A.A. (1947), "The Theory of Enterprise Entity," *Columbia Law Review*, 47 (3): 343–58, reprinted in this volume.

—— (1965), "The Impact of Corporation on Classical Economic Theory," *The Quarterly Journal of Economics*, 79 (1): 25–40, reprinted in this volume.

Berle, A.A. and Means, G. (1932–91), *The Modern corporation & Private Property*. New Brunswick and London: Transaction Publishers.

Biondi, Y. (2005), "The Firm as an Entity: Management, Organization, Accounting," University of Brescia, Dept. Economia Aziendale, Paper 46. Available online at http://ssrn.com/abstract=774764

Blair, M.M. and Stout, L.A. (1999), "A Team Production Theory of Corporate Law," *Virginia Law Review*, 85 (2): 247–328.

Bolton, P. and Scharstein, D.S. (1998), "Corporate Finance, the Theory of the Firm, and Organization," *Journal of Economic Perspectives*, 12 (4): 95–114.

Bowles, S. (1985), "The Production Process in a Competitive Economy: Walrasian, Neo-Hobbesian, and Marxian Models," *The American Economic Review*, 75 (1): 16–36.

Bowles, S. and Gintis, H. (1993), "The Revenge of Homo Economicus: Contested Exchange and the Revival of Political Economy," *Journal of Economic Perspectives*, 7 (1): 83–102.

Chandler, A.D. Jr (1990), *Scale and Scope, The Dynamics of Industrial Capitalism*, Cambridge (Mass.): The Belknap Press of Havard University Press.

—— (1992), "Organizational Capabilities and the Economic History of the Industrial Enterprise," *Journal of Economic Perspectives*, 6 (3): 79–100.

Cheung, S. (1983), "The Contractual Nature of the Firm," *Journal of Law and Economics*, 26: 1–21.

Coase, R.H. (1937), "The Nature of the Firm," *Economica,* 4: 386–405.

Cohen, M.D., Burkhart, R., Dosi, G., Egidi, M., Marengo, L., Warglien, M. and Winter, S. (1996), "Routines and Other Recurring Action Patterns of Organizations: Contemporary Research Issues," *Industrialand Corporate Change*, 5: 653–98.

Cohendet, P. (1998), "*Information, connaissances et théorie de la firme évolutionniste*" [Information, knowledge, and evolutionary theory of the firm], in P. Petit (Sous la direction de), *L'économie de l'information*, Paris: La Découverte.

Commons, J.R. (1924–95), *Legal Foundations of Capitalism*, New edition, New Brunswick: Transaction Publishers.

—— (1932–90), *Institutional Economics*, 2 volumes, New edition, New Brunswick: Transaction Publishers.

Coriat B. and Dosi, G. (1998), "Learning how to Govern and Learning how to Solve Problems: On the Co-Evolution of Competencies, Conflicts and Organizational Routines," in A.D. Chandler Jr, P. Hagström and Ö. Sölvell (eds), *The Dynamic Firm*, Oxford: Oxford University Press.

Coriat, B. and Weinstein, O. (1995), *Les nouvelles théories de l'entreprise* [The new theories of the firm], Paris: Le livre de poche, Librairie Générale Française.

—— (1998), " Sur la théorie évolutionniste de la firme " [On the evolutionary theory of the firm], in R. Delorme, Lemoigne, J.-L., Basle, M. and Paulré, B. (eds), *Approches évolutionnistes de la firme et de l'industrie – théories et analyses empiriques,* Paris: L'Harmattan.

Demsetz, H. (1967), "Toward a Theory of Property Rights," *American Economic Review*, 57: 347–9.

—— (1988), *Ownership, Control, and the Firm*, Oxford: Basil Blackwell.

Dosi, G., Faillo, M. and Marengo, L. (2003), "Organizational Capabilities, Patterns of Knowledge Accumulation and Governance Structure in Business Firms", LEM Working Papers Series, 2003/11. Sant'Anna School of Advanced Studies.

Dosi, G., Levinthal, D. and Marengo, L. (2001), "Bringing Contested Terrain: Linking Incentive-Based and Learning Perspectives on Organizational Evolution", LEM Working Papers Series, 2001/20. Sant'Anna School of Advanced Studies.

Dosi, G. and Marengo, L. (1994), "Some elements of an Evolutionnary Theory of Organizational Competencies," in R.W. England (ed.), *Evolutionary Concepts in Contemporary Economics,* Ann Arbor: University of Michigan Press.

Dosi, G., Nelson, R.R. and Winter, S.G. (eds) (2000), *The Nature and Dynamics of Organizational Capabilities,* Oxford: Oxford University Press.

Ellerman, D.P. (1992), *Property & Contract in Economics*, Cambridge (Mass.): Blackwell Publishers.

Eymard-Duvernay, F. (2004), *Economie Politique de l'entreprise* [Political Economy of the Firm], Paris: Editions La Découverte.

Fama, E.F. (1980), "Agency Problems and The Theory of the Firm," *Journal of Political Economy*, 88 (2): 163–74.

Fama, E.F. and Jensen, M.C. (1983), "Separation of Ownership and Control," *Journal of Law and Economics*, 26: 301–26.

Foss, N.J. (1997), *Resources, Firms, and Strategies: A Reader in the Resource-Based Perspective.* Oxford: Oxford University Press.

Furubotn E.G. and Pejovich, S. (1972), "Property Rights and Economic theory: A Survey of Recent Literature," *Journal of Economic Literature*, 10: 1137–62.

Gibbons, R. (2005), "Four formal(izable) theories of the firm," *Journal of Economic Behavior & Organization*, 58: 200–45.

Grossman, S. and Hart, O. (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94 (2): 691–719.

Hart, O. (1989), "An Economist's Perspective on the Theory of the Firm?," *Colombia Law Review*, 89: 1757–74.

—— (1991), "Incomplete Contracts and the Theory of the Firm," in O.E. Williamson and S.G. Winter (eds), *The Nature of the Firm*, Oxford: Oxford University Press.

—— (1995), *Firms, Contracts, and Financial Structure*, New York: Oxford University Press

Hart, O. and Moore J. (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98 (6): 1119–58.

Hayek, F.A. (1945), "The Use of Knowledge in Society," *American Economic Review*, 35 (4).

Hodgson, G.M. (1998), "Competencies and contracts in the theory of the firm," *Journal of Economic Behavior & Organization*, 35: 179–201.

—— (1999), *Evolution and institutions: on evolutionary economics and the evolution of economics*, Northampton, MA: Edward Elgar Publishing.

Holmstrom, B. (1999), "The Firm as a Subeconomy," *Journal of Law, Economics and Organization*, 15 (1): 74–101.

Holmstrom**,** B. and Roberts R. (1998), "The Boundaries of the Firm Revisited," *Journal of Economic Perspectives*, 12 (4): 73–94.

Jensen, M.C. and Meckling, W. (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Capital Structure," *Journal of Financial Economics*, 3: 305–60.

— (1992–98), "Specific and General Knowledge and Organizational Structure," in L. Werin and H. Wijkander (eds), *Contract Economics*, Oxford: Blackwell; reprints in M.C. Jensen, *Foundations of Organizational Strategy*, Cambridge (Mass.): Harvard University Press.

Klein, B., Crawford, R. and Alchian, A.A. (1978), "Vertical Integration, Appropriable Rents, and the Competitive Contracting Process," *Journal of Law andEconomics*, 21 (2): 297–326.

Kreps, D.M., (1996), "Markets and hierarchies and (mathematical) economic theory," *Industrial and Corporate Change* 5: 561–97.

Langlois, R.N. and Robertson, P.L. (1995), *Firms, Markets and Economic Change. Adynamic Theory of Business Institutions*, London and New York: Routledge.

Leonard, D.H. (1998), *Wellspring of Knowledge*, Boston (Mass.): Harvard Business School Press.

Marengo, L. and Dosi, G. (2005), "Division of labor, organizational coordination and market mechanisms in collective problem-solving," *Journal of Economic Behavior & Organization,* 58: 303–26

Marengo, L., Dosi, G, Legrenzi, P. and Pasquali, C. (1999), "The Structure of Problem-Solving Knowledge and the Structure of Organizations," prepared for the conference "The Roots and Branches of Organizational Economics," SCANCOR, Stanford University, September.

Marshak, J. and Radner, R. (1972), *Economic Theory of Teams*, New Haven: Yale University Press.

Masten, S.E. (1988), "A Legal Basis for the Firm," *Journal of Law, Economics and Organization,* 4 (1): 181–98.

— (2002), "Modern evidence on the firm" *American Economic Review*, 92: 428–32.

Montgomery, C.A. (ed.) (1995), *Resource-based and Evolutionary Theories of the Firm*, Boston/Dordrecht/London: Kluwer Academic Publishers.

Nelson, R.R. and Winter, S.G. (1982, *An Evolutionary Theory of Economic Change.* Cambridge (Mass.): Belknap Press of Harvard University Press.

Nelson, R.R., Winter, S G. and Dosi, G. (2000), *The nature and dynamics of organizational capabilities*. New York: Oxford University Press.

Nonaka, I. (1994), "A Dynamic Theory of Organizational Knowledge Creation," *Organizational Science*, 5 (1): 14–37.

Penrose, E.T. (1959), *The Theory of the Growth of the Firm*, Oxford: Basil Blackwell.

Putterman, L. and Kroszner, R.S. (1996), "The economic nature of the firm: a new introduction," in L. Putterman and R.S. Kroszner (eds.), *The economic nature of the firm. Areader*, Cambridge: Cambridge University Press.

Radner, R. (1986), "The Internal Economy of large Firms," *The Economic Journal*, 96, Supplement: Conference Papers: 1–22.

—— (1992), "Hierarchy: The Economics of Managing," *Journal of Economic Literature*, 30 (3): 1382–415.

Rajan, R.G. and Zingales, L. (1998), "Power in a Theory of the Firm," *Quarterly Journal of Economics*, 113 (2): 387–432.

—— (2000), "The Governance of the New Enterprise," in Xavier Vives (ed.), *Corporate Governance: Theoretical and Empirical Perspectives*, Cambridge: Cambridge University Press.

—— (2001a), "The Influence of the Financial Revolution on the Nature of Firms," *American Economic Review*, 91 (2), 203–11

—— (2001b), "The Firm as a Dedicated Hierarchy: A Theory of the Origins and the Growth of Firms," *Quarterly Journal of Economics*, 116 (3), 805–51

Rumelt, R.P. (1984), "Towards a Strategic Theory of the Firm," in R. B. Lamb (ed.), *Competitive Strategic Management*, Englewood Cliffs, NJ: Prentice-Hall.

Simon, H.A. (1951), "A formal theory of employment relationship," *Econometrica* 19: 293–305.

—— (1991), "Organizations and Markets," *Journal of Economic Perspectives*, 5 (2): 25–44, reprinted in this volume.

—— (1996), *The Sciences of the Artificial*, third edition, Cambridge (Mass.): MIT Press.

—— (1997), *An Empirically based Microeconomics*, Cambridge: Cambridge University Press.

Teece, D.J.G. and Pisano, G. (1994), "The Dynamic Capabilities of Firms: An Introduction," *Industrial and Corporate Change*, 3 (3): 537–56.

Teece, D. J.G., Pisano, G. and Shuen, A. (1997), "Dynamic Capabilities and Strategic Management," *Strategic Management Journal*, 18 (7): 509–33.

Tirole, J. (1999), "Incomplete Contracts: Where Do We Stand?," *Econometrica*, 67 (4): 741–81.

Weinstein O. and Azoulay, N. (1999). "Firms' Capabilities and Organizational Learning. A critical survey of some literature," paper prepared for the DYNACOM research project, funded by Targeted Socio-Economic Research (TSER) Area 1.1, under the Fourth Framework Programme, European Commission.

Wernerfelt, B. (1984), "A Resource-Based View of the Firm," *Strategic Management Journal*, 5 (2): 171–80.

Williamson, O.E. (1979), "Transaction cost economics: the governance of contractual relations," *Journal of Law and Economics*, 22: 233–61.

—— (1985), *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*, New York: Free Press.

—— (1991), "The Logic of Economic Organization," in O.E Williamson and S.G. Winter (eds), *The Nature of the Firm*, Oxford: Oxford University Press.

—— (2000), "The New Institutional Economics: Taking Stock, Looking Ahead," *Journal of Economic Literature*, 38(2): 595–613.

Winter S.G. (1982), "An essay on the theory of production," in S.H. Hyman (ed.), *Economics and the World Around it.* Ann Arbor: University of Michigan Press.

Zingales, L. (2000), "In Search of New Foundations," *Journal of Finance*, 55 (4): 1623–53.

# 4 Organizations and markets*

*Herbert A. Simon*

In classical and neoclassical *economic* theory, markets are at the center of the stage. The actors in these markets are workers and consumers (sometimes combined into households), firms, owners of resources, governments, and perhaps others. The economic world of the neoclassical textbooks is a world of transactions, and these transactions typically involve an exchange of goods, services, and/or money that both parties to the transaction find advantageous to achieve these goals. Along with consumption, work and leisure are important components of the utility functions of households. Often, profit is assumed to be the sole objective of firms and their owners.

The description of the parties who participate in these transactions is minimal. However, as soon as firms are elaborated to become more than simple nodes in a network of transactions, to be producers – transformers of "factors" into products – difficult and important questions arise for the theory. A large part of the behavior of the system now takes place inside the skins of firms, and does not consist just of market exchanges. Counted by the head, most of the actors in a modern economy are employees, who either do not spend their days in trading, or if they do (for example, if they are salesmen or purchasing agents) are assumed to trade as agents of the firm rather than in their own interest, which might be quite different.

This raises the question of why there are firms at all. Why are not all the actors independent contractors? Why do most of them enter into employment contracts, selling their labor for a wage? What determines the make-or-buy decisions of firms, hence the boundaries between them and markets? When will two domains of activity lie within a single firm, and when will they be handled by separate contracting firms?

A second set of questions asks how the employees of firms are motivated to work for the maximization of the firm's profit. What's in it for them? How are their utility functions reconciled with those of the firm? In the employee's utility function, work is usually assumed to have negative utility and leisure (including loafing and working lackadaisically) to have positive utility. Why do employees often work hard?

The simple (neoclassical) answer to the motivational question derives from the employment contract, under which workers maximize their utility by accepting

the authority of the firm; that is, by agreeing to accept orders from the profit maximizers in charge. But this answer leads to the new question of how the employment contract is enforced by the employer. In particular, how are employees induced to work more than minimally, and perhaps even with initiative and enthusiasm? Why should employees attempt to maximize the profits of their firms when making the decisions that are delegated to them?

These questions about the scope of activity and operation of firms have spawned a vigorous cottage industry, a branch of which is sometimes called "the new institutional economics," which tries to explain when activities will be carried out through the market and when they will be carried out within the skins of firms, and tries to explain also how it is possible for firms to operate efficiently. In the literature of the new institutional economics, two ideas that play a major role in the explanations are "transaction costs" and "opportunism" (for example, Williamson (1975), (1985)). Sometimes the explanations are couched in terms of "information asymmetry" or "incomplete information" (Ross (1973); Stiglitz (1974)). In other writings these topics are subsumed under agency theory, which treats the employment contract as an optimal contract between principal and agents, and studies how contractual arrangements can deal with shirking and other motivational problems.

The idea behind these ideas is that a proper explanation of an economic phenomenon will reduce it to maximizing behavior of parties who are engaged in contracting, given the circumstances that surround the transaction. The terms of the contract will be influenced by the access of the parties to information, by the costs of negotiating, and by the opportunities for cheating. Access to information, negotiation costs, and opportunities for cheating are most often treated as exogenous variables that do not themselves need to be explained. It has been observed that they even introduce a sort of bounded rationality into the behavior, with the exogeneity of the limits of rationality allowing the theory to remain within the magical domains of utility and profit maximization.

A fundamental feature of the new institutional economics is that it retains the centrality of markets and exchanges. All phenomena are to be explained by translating them into (or deriving them from) market transactions based upon negotiated contracts, for example, in which employers become "principals" and employees become "agents." Although the new institutional economics is wholly compatible with and conservative of neoclassical theory, it does greatly multiply the number of auxiliary exogenous assumptions that are needed for the theory to work. For example, to explain the presence or absence of certain kinds of insurance contracts, moral risk is invoked; the incompleteness of contracts is assumed to derive from the fact that information is incomplete or distributed asymmetrically between the parties to the contract. Since such constructs are typically introduced into the analysis in a casual way, with no empirical support except an appeal to introspection and common sense, mechanisms of these sorts have proliferated in the literature, giving it a very *ad hoc* flavor.

In general, the new institutional economics has not drawn heavily from the empirical work in organizations and decision-making for its auxiliary assumptions.

(For introductions to that literature, see March and Simon (1958); Cyert and March (1963): Kornai (1971); Simon (1979)). Nevertheless, it is appropriately subversive of neoclassical theory in that it suggests a whole agenda of microeconomic empirical work that must be performed to estimate the exogenous parameters and to test the theory empirically. Until that research has been carried out (and the existing literature on organizations and decision making taken into account), the new institutional economics and related approaches are acts of faith, or perhaps of piety.

## The ubiquity of organizations

A mythical visitor from Mars, not having been apprised of the centrality of markets and contracts, might find the new institutional economics rather astonishing. Suppose that it (the visitor – I'll avoid the question of its sex) approaches the Earth from space, equipped with a telescope that reveals social structures. The firms reveal themselves, say, as solid green areas with faint interior contours marking out divisions and departments. Market transactions show as red lines connecting firms, forming a network in the spaces between them. Within firms (and perhaps even between them) the approaching visitor also sees pale blue lines, the lines of authority connecting bosses with various levels of workers. As our visitor looked more carefully at the scene beneath, it might see one of the green masses divide, as a firm divested itself of one of its divisions. Or it might see one green object gobble up another. At this distance, the departing golden parachutes would probably not be visible.

No matter whether our visitor approached the United States or the Soviet Union, urban China or the European Community, the greater part of the space below it would be within the green areas, for almost all of the inhabitants would be employees, hence inside the firm boundaries. Organizations would be the dominant feature of the landscape. A message sent back home, describing the scene, would speak of "large green areas interconnected by red lines." It would not likely speak of "a network of red lines connecting green spots."

Of course, if the vehicle hovered over central Africa, or the more rural portions of China or India, the green areas would be much smaller, and there would be large spaces inhabited by the little black dots we know as families and villages. But the red lines would be fainter and sparser in this case, too, because the black dots would be close to self-sufficiency, and only partially immersed in markets. But let us, for the present, restrict our attention to the landscape of the developed economies.

When our visitor came to know that the green masses were organizations and the red lines connecting them were market transactions, it might be surprised to hear the structure called a market economy. "Wouldn't 'organizational economy' be the more appropriate term?" it might ask. The choice of name may matter a great deal. The name can affect the order in which we describe its institutions, and the order of description can affect the theory. In particular, it

may strongly affect our choice of the variables that are important enough to be included in a first-order theory of the phenomena.

How does the economy look when it is viewed as an organizational economy, with market relations among organizations? I have already suggested some of the more prominent features.

First, most producers are employees of firms, not owners. Viewed from the vantage point of classical theory, they have no reason to maximize the profits of firms, except to the extent that they can be controlled by owners. Moreover, profit-making firms, nonprofit organizations, and bureaucratic organizations all have exactly the same problem of inducing their employees to work toward the organizational goals. There is no reason, *a priori*, why it should be easier (or harder) to produce this motivation in organizations aimed at maximizing profits than in organizations with different goals. If it is true in an organizational economy that organizations motivated by profits will be more efficient than other organizations, additional postulates will have to be introduced to account for it.

Second, the system is nearly in neutral equilibrium between the use of market transactions and authority relations to handle any particular matter: that is to say, very small changes in the situation can tip the equilibrium one way or the other. It is hard to explain degrees of integration of economic activities. In many instances, transaction cost analysis is not applicable, and even where it is, there often remains considerable latitude for different degrees of integration. For example, why are auto dealerships not a part of auto manufacturing companies, rather than having contractual relations with them?[1] Why did General Motors manage its own tool design for many years, but recently decide to contract most of it out? Under constant returns to scale and reasonably competitive markets, which characterize many manufacturing situations, make-or-buy decisions become ambiguous. The possibility of using internal division-by-division balance sheets, and internal pricing in negotiation between components of an organization, further blurs the boundary between organizations and markets.

Without the introduction of very particular *ad hoc* assumptions, unbuttressed by empirical evidence, neoclassical theory provides no explanation for the repeated appearance of Pareto distributions of business firm sizes in virtually all situations where size distributions have been studied (Ijiri and Simon (1977); Simon (1979)). (In a Pareto distribution, the logarithm of the number of firms above any given size decreases linearly with the logarithm of the size.) These observed distributions are difficult to reconcile with any notions that have been proposed for optimal firm size, but are easily explained by simple, plausible probabilistic mechanisms that make no appeal to optimality.

In sum, an organizational economy poses the questions of why the larger part of a modern economy's business is done by organizations, what role markets play in connecting these organizations with each other, and what role markets play in connecting organizations with consumers. Moreover, the boundary between markets and organizations varies greatly from one society to another and from one time to another. What mechanism maintains the highly fluid

equilibrium between them? Until these questions are answered, it will be difficult to draw conclusions about the relative efficiencies of different forms of ownership and control of organizations, or the relative efficiency of markets versus central planning.

## Motivation and efficiency in organizations

There are three different questions of social organization that are usually confounded, but which need to be considered separately. The first is the question of the relative efficiency of markets and organizations. The second is the question of the consequences of having a society's organizations owned by profit-making organizations, by nonprofit organizations, or by public organizations, respectively. The third is the question of the consequences of using central planning instead of markets to regulate relations among organizations. At present, our concern is only with the first question: what makes organizations work as well or badly as they do?

In particular, for whom is profit the motive? Adolf Berle and Gardiner Means posed the problem very sharply in their famous book, *The Modern Corporation and Private Property* (1933), by showing that even at the top executive levels of the modern corporation there is a great gap between ownership and control, and a correspondingly great opportunity for discrepancy between the goal of owners (profit) and the goals of managers (career status, wealth, a quiet life, and so on).

Demsetz and Lehn (1985) have contested the argument of Berle and Means on the ground that even large corporations show considerable concentration of ownership. Typically, a half dozen owners (or fewer) own 10 or 20 percent of the shares, enough to retain controlling power. Often, these owners are also the active top executives. But the objection does not hold water. If a company has an executive bonus plan, and if an executive's percentage share in bonus awards is greater than his or her percentage share of dividends, then it pays that executive to divert earnings from dividends to bonuses. Most companies have executive award systems that make this conflict of interest very real. Golden parachutes and leveraged buyouts are other significant examples of transactions where the interests of shareholders and executives may diverge strongly.[2]

If even top executives may be conflicted in their motives, the problem should be still greater for employees who are not owners at all, or only insignificantly. Principal–agent theory, on which the new institutional economists often rely, assumes that agents within firms will shirk unless their actions contribute directly to their own economic self-interest. It is only via monitoring combined with contracts that appeal to their self-seeking nature that such shirking may be mitigated. But the assumption that executives (and perhaps other employees) would choose to advance their own careers and wealth and consumption, rather than pursuing organizational goals like maximizing profit, is not prescribed by neoclassical theory, which leaves the specification of the utility function completely open.

Why not assume that maximizing the firm's profit is precisely what maximizes the utilities of executives and other workers? In a society of robots, an owner would not settle for less. But most of us would think this an unrealistic assumption to make for a human society. An organization theory with an unspecified utility function is not a theory at all. And one with an unrealistic utility function does not provide a basis for understanding real organizations. Instead, we should begin with empirically valid postulates about what motivates real people in real organizations. I shall argue that such postulates can be derived from four organizational phenomena whose roles are amply documented in the literature on organizations: authority, rewards, identification, and coordination.

## Authority: the employment relation

The employment contract is an example of what is now sometimes called an "incomplete contract;" that is to say, some of its terms are unspecified. Employees agree to do, over the life of the contract, what they are ordered to do; but the orders will not be issued until some time after the contract is negotiated (Simon (1951); Williamson (1975)).

The usual argument (within the neoclassical framework) for the existence of incomplete contracts is that in a world of uncertainty actions will have to be taken as the situation calls for them, without time for negotiation. The employee is rewarded, in the level of the wage, for willingness to bear the brunt of this uncertainty as to what actions will be chosen, and to do, when the time comes, whatever the employer thinks the situation calls for. This argument does not imply that uncertainty is replaced by complete certainty at the time of decision. On the contrary, taking decisions under conditions of uncertainty may be one of the important skills demanded of the decision maker. The essential point is that the uncertainty for the employer is decreased by delaying the commitment to specific actions from the time employment begins until the time when action is called for.

An employment contract contains all sorts of implicit (and explicit) limitations that set boundaries to the range of actions the employee will be directed to perform. These boundaries define the "zone of acceptance" within which an employee can be expected to obey orders. The zone of acceptance is also sometimes called a "zone of indifference," for the choice among alternative behaviors, while of major importance to the employer, may be of little or no concern to the employee. A secretary, for example, usually has little or no preference for typing a letter to one of the company's customers rather than another, and little interest in the content of the letter. Even a factory manager will accept, within wide limits, whatever mix of products the factory is ordered to produce in a given month.

The combination of uncertainty on the part of the employer (as to what will need to be done in the future) and broad acceptance of the employee (of what he or she will be ordered to do) makes the employment contract a very attractive bargain for both parties. The new institutional economics finds that

employment achieves great savings in transaction costs – the costs of negotiating separate contracts for each action.

But this theory of the employment contract must be elaborated. Authority in organizations is not used exclusively, or even mainly, to command specific actions. Most often, the command takes the form of a result to be produced ("repair this hinge"), or a principle to be applied ("all purchases must be made through the purchasing department"), or goal constraints ("manufacture as cheaply as possible consistent with quality"). Only the end goal has been supplied by the command, not the method of reaching it. The mechanic must apply all kinds of knowledge and skill to repairing the hinge. The section chief must initiate purchases of supplies needed for the work of that section; however, the company's standard procedures must be taken as ground rules for the way the purchases are made. The factory manager must control manufacturing cost and quality.

Employees, especially but not exclusively at managerial and executive levels, are responsible not only for evaluating alternatives and choosing among them but also for recognizing the need for decisions, putting them on the agenda, and seeing to the generation of possible actions. Doing the job well is not mainly a matter of responding to commands, but is much more a matter of taking initiative to advance organizational objectives.

Commands do not usually specify concrete actions but, instead, define some of the premises that are to be used by employees in making the decisions for which they are responsible (Simon (1947)). Hence, seeing that commands are obeyed is not simply a matter of observing behavior, but of affecting the thought processes and the decision premises of employees. Further, it is usually difficult or impossible to ascertain what these decision premises have been without reviewing the whole decision – thus causing an almost complete loss of the economy that was sought in delegating it in the first place.

The command an employer might like to issue is: "Always decide in such a way as to maximize company profit!" But that would simply reintroduce the question of how the extent of obedience to the command is to be observed without losing the benefit of delegation. Even if the employees were robots, whose loyalty could be guaranteed, the problem would not be solved. For giving each robot complete discretion would surrender large efficiencies usually attainable from specialization in decision-making work. We need to delegate within guidelines, which creates the problem of monitoring the observance of guidelines without recentralizing what has just been delegated.

If authority is used to transmit premises for making decisions rather than commands for specific behaviors, then many different experts can contribute their knowledge to a single decision. Information and policy rules can flow through the organization along many channels, serving as inputs – decision premises – for many organizational behaviors.

The accounting department gathers cost data, which it supplies to the head of the blast furnace department to help make operating decisions in that department. At the same time, the blast furnace manager is receiving instructions from

metallurgical specialists on the technical aspects of the operation. The faint blue lines that our visitor from Mars saw within the green areas were not just streams of orders, but flows of all kinds of decision premises (constraints and information as well as orders) from one point in the organization to another.

This explication of the employment contract and authority takes us back to the question of motivation. For the organization to work well, it is not enough for employees to accept commands literally. In fact, obeying operating rules literally is a favorite method of work slowdown during labor–management disputes, as visitors to airports when controllers are unhappy can attest. What is required is that employees take initiative and apply all their skill and knowledge to advance the achievement of the organization's objectives.

We should not assume without evidence that organizations do work well. But "well" is a relative term. In most organizations, employees contribute much more to goal achievement than the minimum that could be extracted from them by supervisory enforcement of the (vague) terms of the employment contract. Why do employees not substitute leisure for work more consistently than they do? Why do they often work so vigorously for the welfare of the organization?

## Rewards as motivations

One obvious answer to the motivational question is that employees may be motivated to accept authority by giving them material rewards, promotion, and recognition for advancing the organization's goals as defined by management. Such rewards certainly provide motivation, but they only operate satisfactorily when certain conditions are met. The most important condition is that the employee's contribution to the organization's goals must be measurable with reasonable accuracy. For example, salesmen are frequently compensated (at least partly) on a commission basis. Blue-collar employees are sometimes compensated on a piecework basis, albeit in a continually decreasing number of situations. Executives, and sometimes others, receive bonuses that are supposed to be related to their contributions to profits.

But such reward systems are effective only to the extent that success can be attributed accurately to individual behaviors. If the indices used to measure outcomes are inappropriate, either because they do not measure the right variables, or because they do not properly identify individual contributions, then reward systems can be grossly inefficient or even counterproductive. Where output quantities are measured with inadequate attention to quality, response to rewards will cause quantities to grow at the cost of lowered quality. Where compliance with company policies that constrain action is not measured, constraints will be ignored and violated. Salesmen may misrepresent the product, workmen may ignore safety rules, managers may buck difficulties to other departments.

In general, the greater the interdependence among various members of the organization, the more difficult it is to measure their separate contributions to the achievement of the organizational goals. But of course, intense interdependence

is precisely what makes it advantageous to organize people instead of depending wholly on market transactions. The measurement difficulties associated with tying rewards to contributions are not superficial, but arise from the very nature and rationale of organization.

Many large U.S. corporations attempted to respond to this problem in the years after World War II by slicing their organization into components that were relatively self-contained. Then, separate balance sheets could be maintained for each division, and these balance sheets could be used to evaluate results and to compute rewards.

Of course, divisionalization can be successful only to the extent that the divisions are actually self-contained. If one division operates mainly as a supplier of parts to other divisions, then policies have to be laid down for setting the prices for items "sold" by the one division to the others, and for determining under what conditions a division may go outside the company to purchase items at a lower price. For these and similar reasons, divisionalization can only be carried a short distance down the structure of a typical corporation, and solves the problem of attributing outcomes to individuals only at the higher levels, if at all.

Although economic rewards play an important part in securing adherence to organizational goals and management authority, they are limited in their effectiveness.[3] Organizations would be far less effective systems than they actually are if such rewards were the only means, or even the principal means, of motivation available. In fact, observation of behavior in organizations reveals other powerful motivations that induce employees to accept organizational goals and authority as bases for their actions. We turn next to the most important of these mechanisms: organizational identification.

## Loyalty: identification with organizational goals

Pride in work and organizational loyalty are widespread phenomena in organizations (Simon (1947)). These traits are more strongly evident among skilled and managerial employees than among employees engaged in very routine work. (The latter are also more easily supervised, and can sometimes be rewarded on a piecework basis.) In part, these attitudes can be attributed to the linkage between an organization's overall success and the personal careers and monetary rewards it can provide its employees. But this explanation ignores the problem of the commons – of benefits that are jointly gained and shared by all, non-contributors along with contributors – and the consequent possibilities for free-riding. The quality and success of an organization depends very little on the energy of any single employee (except possibly an executive at or near the very top). Why will employees work hard if they can gain almost as much by loafing?

Of course free-riding can be observed in organizations. The elimination of free riding is generally thought to be the principal reason for the success of the Chinese agricultural reforms after 1980, when responsibility and reward for agricultural production were transferred from the commune to the family. The question is not whether free riders exist – much less employees who exert

something less than their maximum – but why there is anything besides free-riding. Why do many workers, perhaps most, exert more than minimally enforceable effort? Why do employees identify with organizational goals at all?

Contemporary evolutionary theory has cautioned us against postulating altruistic motives for people. In models of natural selection, nice guys generally aren't fit; they don't multiply as rapidly as their more selfish brethren. The argument from natural selection has often been used, explicitly or implicitly, to fill the utility function with selfish personal goals. But models of natural selection do not actually provide strong support for the idea that people will only pursue selfish personal economic goals. In fact, such models in no way foreclose the possibility (indeed, the probability) that people will be strongly motivated by organizational loyalty, even when they can expect no "selfish" rewards from it (Simon (1983); (1990)).

First, it should be emphasized that what natural selection increases is fitness, the number of progeny of the successful competitor. But in modern society, the attainment of wealth or other selfish rewards is not directly connected to number of progeny. In fact, first-world societies generally display a negative correlation between income level and size of family. But let us waive this point, as distracting us from the main argument, and suppose that attainment of the goals usually described as selfish (especially personal economic goals) contributes to evolutionary fitness.

We come then to the second point: each human being depends for survival on the immediate and broader surrounding society. Human beings are not the independent windowless Leibnitzian monads sometimes conjured up by libertarian theory. Society is not imposed on humans; rather, it provides the matrix in which we survive and mature and act on the environment. Families and the rest of society provide nutrition, shelter, and safety during childhood and youth, and then the knowledge and skills for adult performance. Moreover, society can react to a person's activities at every stage of life, either facilitating them or severely impeding them. Society has enormous powers, enduring though a person's lifetime, to enhance or reduce evolutionary fitness.

What kinds of traits, in addition to personal strength and intelligence, would contribute to the fitness of this socially dependent creature? One such trait, or combination of traits, might be called docility. To be docile is to be tractable, manageable, and above all, teachable. Docile people tend to adapt their behavior to norms and pressures of the society. I am not satisfied that "docile" conveys my meaning precisely, but I know of no better word.

That fitness is derivable from being docile becomes evident when we consider the opposite of docility: intractability, unmanageability, unteachability, incorrigibility. The argument is not that people are totally docile, nor that they are totally selfish, but that fitness calls for a measured but substantial responsiveness to social influence. In some contexts, this responsiveness implies motivation to learn or imitate; in other contexts, willingness to obey or conform. From an evolutionary standpoint, having a considerable measure of docility is not altruism but enlightened selfishness.

To survive as a trait, docility must contribute on average to the fitness of the individual who possesses it. Yet it may still lead to self-injurious behavior in particular cases. Thus, docile individuals may do better at earning a living, but loyalty to the nation may lead them to sacrifice their lives in wartime. Once docility is present, society may exploit it by teaching values that are truly altruistic; that is, which contribute to the society's fitness, but not to the individual's. The only requirement is that on balance and on the average the docile individual must be fitter than the one who is not docile.[4]

Of course, showing that a configuration of traits or genes would contribute to fitness, if they existed, does not prove they exist. But ample empirical evidence shows that most human beings are gifted with a considerable measure of docility. The purpose of the present argument is to show that this docility and the altruism it induces is wholly consistent with the premise of selection of the fittest. In fact, the theory of natural selection strongly predicts the appearance of docility and altruism in social animals.

Docility is used to inculcate individuals with organizational pride and loyalty. These motives are based upon a discrimination between a "we" and a "they." Identification with the "we," which may be a family, a company, a city, a nation, or the local baseball team, allows individuals to experience satisfactions (to gain utility) from successes of the unit thus selected. Thus, organizational identification becomes a motivation for employees to work actively for organizational goals. Of course, identification is not an exclusive source of motivation; it exists side by side with material rewards and enforcement mechanisms that are part of the employment contract. But a realistic picture of how organizations operate must include the importance of identification in the motivations of employees.

The strength of organizational identifications will depend upon the extent to which a society uses the docility mechanism to inculcate them, and this appears to vary considerably from one society to another. For instance, it would probably be agreed by ethnographers that in Chinese society greater pressure is exerted to induce identification with the family than with employing organizations, while the reverse is true of Japanese society. Such conjectures can be tested, for example, by examining practices of nepotism, and attitudes toward it, in the two societies.

The strength of the organizational loyalties of employees is not to be attributed only to motivation induced by docility. There is also an important cognitive component. The bounded rationality of humans does not allow us to grasp the complex situations that provide the environments for our actions in their entirety. The first step in rational action is to focus attention on specific (strategic) aspects of the total situation, and to form a model of the situation in terms of those aspects that lie in that focus of attention. Rational computation takes place in the context of this model, rather than in the response to the whole external reality.

One dimension of simplification is to focus on particular goals, and one form of focus is to attend to the goals of an organization or organization unit. Having defined that unit as the "we," actions are evaluated in terms of their contribution to the unit's objectives. The ubiquity of this narrowing of attention

is easily demonstrated. As one example, Dearborn and Simon (1958) presented a group of business executives with a description of the current situation of a large company, and asked them to identify the most serious problem facing the company. In their own companies, some of the executives were responsible for manufacturing, others for sales, others for finance. In almost every case, the "most serious problem" identified by the executive lay in the domain of his or her own department – manufacturing problems for manufacturing executives, sales problems for sales executives, and so on.

It is a commonplace of organizational life that a person's organizational identification will shift with his or her position, although the motivational basis for the shift is perhaps more widely recognized than the cognitive basis. But a shift in organizational position exposes the employees to new "facts" and phenomena, to a new network of communications, and to new goals. A different model is inevitably formed of the decision-making situation, a model that emphasizes local components of the environment and local goals. Behavior is very much a function of position.

Because of cognitive limits, the precise form that goals take may depend on what can be measured in the situation. In business organizations, the accounting statements provide stylized measurements of profits, size, growth, market share, and so on. Even if these measurements are only rough approximations of the things they are supposed to be measuring, they are likely to replace the "real" unmeasured concepts in the decision-making process.

Willingness of employees at all levels to assume responsibility for producing results – not simply "following the rules" – is generally believed to be a major determinant of organizational success. This discussion implies that acceptance of responsibility will be affected both by the reward system and by the strengths of organizational identifications. Here again, large intercultural differences may exist. The recent establishment of a substantial number of international joint ventures, with managements and employees recruited from different cultures, provides an excellent research environment for studying these differences and their effects upon organizational efficiency.

Since the developments are quite new, little information is yet available about them. But one example where data are available is the joint venture between Toyota and General Motors in northern California (Krafcik and Womack (1987)). Here Toyota took over a former General Motors plant, equipped it with standard state-of-the-art machinery, rehired employees mainly from the previous work force and accepted the same union. They have been able to produce automobiles with about 45 percent fewer labor hours than an entirely comparable GM plant that uses American managers and management methods, and about 30 percent fewer hours than a new GM plant having more modern "hitech" equipment, and only about 15 percent more labor hours than a comparable Toyota plant in Japan.

The causes for these enormous differences in efficiency have almost nothing to do with the classical physical production function. They also appear to have little to do with cultural differences at the blue-collar level.[5] They seem to have

nothing to do, either, with material reward structures, which are not significantly different in the various plants. They must be attributed in large part to differences in management practices (for example, quality control practices, and inventory policies), perhaps bolstered by differences in management attitudes and motivations.

## Coordination

This examination of authority and organizational identification should help explain how organizations can be highly productive even though the relation between their goals and the material rewards received by employees, if it exists at all, is extremely indirect and tenuous. In particular, it helps explain why careful comparative studies have generally found it hard to identify systematic differences in productivity and efficiency between profit-making, nonprofit, and publicly-controlled organizations (Weisbrod (1988), (1989)). Also, it explains why Demsetz and Lehn (1985) found no difference in profits between corporations that were managed or controlled by owners and those with diffuse stock ownership.

But to understand the relative advantages of organizations and markets, and the circumstances under which one would operate more effectively than the other, one further component must be added to our description of organizations. Organizations, through the authority mechanism, provide a means for coordinating the activities of groups of individuals in ways that are not always easily achieved by markets.

Coordination is a rather slovenly word, often abused in organizations. An experienced executive cringes when he or she learns that someone has been appointed to "coordinate" a set of activities, since calling for coordination without specifying just what it means is simply a lazy way of passing off problems to someone else. I will try to make the concept more precise by using it to designate a specific kind of activity.

The theory of games has sharply underscored that decisions are usually indeterminate when each party in a situation is uncertain about the actions of the others. This result is quite independent of whether their goals are complementary or competitive. One simple example of this indeterminacy is that it is rational for a motorist to drive on the same side of the road as other drivers headed the same direction, whichever that may be. There is no question of correct behavior in relation to the environment, but only of coordinating the behaviors of all the actors. Such rules of the road, or standardization, can greatly improve the performance of systems in those (ubiquitous) situations where the correctness of an action depends on what the other actors are doing.

A more complex example of coordination is provided by a university. Conceive of a university that consisted only of some rooms, some teachers, and some students. Students and teachers would "simply" negotiate to meet at certain times and places for their classes. The resulting chaos would probably be resolved by inventing the institutions of a registrar's office and a class schedule. While it would be extravagant to urge that class schedules provide the *raison d'etre* for education by universities, rather than by contractual tutoring arrangements

negotiated through markets, nevertheless, the coordinating function of schedules is not trivial.

A major use of authority in organizations is to coordinate behavior by promulgating standards and rules of the road, thus allowing actors to form more stable expectations about the behavior of the environment (including the behavior of other actors). Since organizations provide a mechanism (authority) for establishing rules of the road, which markets do not, one might even expect organizations rather than markets to be the environments in which the behavior called "rational expectations" would be most often observed.[6]

In a book on central planning during World War II, Ely Devons (1950) raised the question of why prices are supplanted by government plans, expressed as quantity goals for production and allocation, as coordinating mechanisms during wartime. The usual argument for markets, as in the well-known 1945 paper of von Hayek, is that they simplify the decision process by reducing the need of each actor to know what the other actors are doing or what situations confront them. To the extent that markets and prices perform this simplifying function, we would expect them to replace centralized decisions when a situation becomes more complex – for example, during the rapid changes that take place in shifting from a peacetime to a wartime economy. Yet, as Devons points out, it is just at such times that central planning tends to increase. Is this irrationality, or are there valid reasons for the shift?

The answer is rather obvious. Prices perform their informational function when they are known or reasonably predictable. Uncertain prices produced by unpredictable shifts in a system reduce the ability of actors to respond rationally. This point is often made by economists in arguing the costs of unexpected inflation, but its implication for the choice between organizations and markets is less often noted. Nor is it often noted that many kinds of uncertainties other than price uncertainties may make coordination through organizational procedures advantageous.

The difficulty that economics has had in giving a good account of organizations and their predominance is traceable in no small part to the fascination of economists with systems in equilibrium. Analysis under assumptions of perfect knowledge and certain expectations has little relevance, surely, for such issues of economic organization as explaining how an economy is structured between organizations and markets. Prices provide only one of the mechanisms for coordination of behavior, either between organizations or within them. Coordination by adjustment of quantities is probably a far more important mechanism from a day-to-day standpoint, and in many circumstances will do a better job of allocation than coordination by prices. For example, inventory control systems record the amounts of inputs for the organization's activities, and place orders when quantities fall below specified levels. The orders, recorded by the control systems of suppliers, initiate the scheduling of new production and are used to adjust aggregate production levels as well.

From a conceptual standpoint, it is entirely feasible to construct economies in which prices are based on costs and final demands are limited wholly by budget constraints, with demand vectors that are otherwise insensitive to prices. Quantities

of goods sold and inventories, not prices, provide the information for coordinating these systems. The Leontief input–output models, with exogenous vectors of final demands, are examples, and the Hawkins–Simon theorem (1949) states the conditions under which such systems have non-negative solutions. They possess the same information-conserving virtues as price-regulated systems (von Hayek (1945)). Each actor need only know his or her own business.

Many observers of business scheduling and pricing practices have claimed that (with the possible exception of the agricultural and mining sectors) models that use quantities as signals approximate first-world national economies more closely than do models in which prices are the principal mechanisms of coordination. I don't wish to argue that point here: but simply observe that quantity adjustments play a very large role in the real world in equilibrating the operations of different organizations and different parts of organizations.

The stylized market exchanges of neoclassical economic theory generally involve only prices and quantities, which is the foundation for their parsimony in information. But actual contracts negotiated between business firms – putting consumer products aside, for the moment – usually specify far more than prices and quantities. Contracts for construction of a building or of a product of engineering (like a generator or an airplane) specify in enormous detail the specifics of the product to be delivered. They require a massive exchange of information in both negotiation and execution. The red market traces that our Martian visitors observed from space are not narrow tracks along which only money and goods flow, but broad highways to accommodate a vast flow of detailed information as well.

Thus, the assertion that markets permit each firm to do its business with little knowledge of its partners is a fiction. In construction, in heavy industry, in manufacturing involving high technology, and in other areas, contracting partners carry on communication at a level comparable to the levels observed between departments of a firm. When products are manufactured to specifications, a great deal of information must flow among the various groups of people involved in the manufacture. But the widespread use of subcontracting in the automobile and construction industries, just to mention two, demonstrates that it is often quite feasible to transport this information across organizational boundaries, so that vertical integration is unnecessary. From this perspective, the distinction between market communications and internal communications, and the criteria for choosing between the two alternative arrangements, become correspondingly vague.

The choice between prices or quantities to coordinate the activity levels of different organizations or parts of organizations does not by itself dictate the respective roles of organizations and markets. Prices may be used to coordinate the activities of different parts of single organizations, provided that some way can be found to determine what the market prices should be, and quantity adjustments can be made between different organizations as well as within them.

There is one important difference in the operation of coordination mechanisms within and between organizations. Coordination between organizations

depends almost wholly on economic motivations and rewards, and becomes seriously imperfect wherever major externalities are present that cannot be removed by enforceable contract arrangements. Within organizations, on the other hand, identification is a powerful force for combating externalities produced by attachment to subgoals, by virtue of the loyalty it can produce to the goals of the whole system. A department will be less likely to skimp on quality to cut costs if its members identify with the final product. In particular, identification becomes an important means for removing or reducing those inefficiencies that are labelled by the terms "moral hazard" and "opportunism."

These observations nudge us toward the conclusion that organization size and degree of integration, and the boundaries between organizations and markets, are determined by rather subtle forces. The wide range of organizational arrangements observable in the world suggests that the equilibrium between these two alternatives may often be almost neutral, with the level highly contingent on a system's history. A traditional arrangement may be preserved until its inefficiencies become overwhelming – or even beyond. The same conclusion is suggested by the constant flux of mergers and spinoffs in the business world, many of these transformations being governed by considerations quite unrelated to productive or allocative efficiency, and many having consequences for efficiency that even those involved in them cannot evaluate.

Over a span of years, a large fraction of all economic activity has been gathered within the walls of large and steadily growing organizations. The green areas observed by our Martian have grown steadily. Ijiri and I have suggested that the growth of organizations may have only a little to do with efficiency (especially since, in most large-scale enterprises, economies and diseconomies of scale are quite small), but may be produced mainly by simple stochastic growth mechanisms (Ijiri and Simon (1977)).

But if particular coordination mechanisms do not determine exactly where the boundaries between organizations and markets will lie, the existence and effectiveness of large organizations does depend on some adequate set of powerful coordinating mechanisms being available. These means of coordination in organizations, taken in combination with the motivational mechanisms discussed earlier, create possibilities for enhancing productivity and efficiency through the division of labor and specialization.

In general, as specialization of tasks proceeds, the interdependency of the specialized parts increases. Hence a structure with effective mechanisms for coordination can carry specialization further than a structure lacking these mechanisms. It has sometimes been argued that specialization of work in modern industry proceeded quite independently of the rise of the factory system. This may have been true of the early phases of the industrial revolution, but would be hard to sustain in relation to contemporary factories. With the combination of authority relations, their motivational foundations, a repertory of coordinative mechanisms, and the division of labor, we arrive at the large hierarchical organizations that are so characteristic of modern life.

## Conclusions

The economies of modern industrialized society can more appropriately be labelled organizational economies than market economies. Thus, even market-driven capitalist economies need a theory of organizations as much as they need a theory of markets. The attempts of the new institutional economics to explain organizational behavior solely in terms of agency, asymmetric information, transaction costs, opportunism, and other concepts drawn from neoclassical economics ignore key organizational mechanisms like authority, identification, and coordination, and hence are seriously incomplete.

The theory presented here is simple and coherent, resting on only a few mechanisms that are causally linked. Better yet, it agrees with empirical observations of organizational phenomena. Large organizations, especially governmental ones, are often caricatured as "bureaucracies," but they are often highly effective systems, despite the fact that the profit motive can penetrate these vast structures only by indirect means.

This theory of organizations calls for reexamining some of the classical questions of political economy. The primacy of profit as the enforcer of organizational efficiency is replaced by organizational goals, combined with organizational identifications and with material rewards and supervision, all of which motivate employees to work toward these goals. This framework makes it necessary to reopen the question of when profit-making, nonprofit, and governmental organizations should be expected to operate well, and when market competition is needed to discipline organizations to perform efficiently.

The reopening of these questions is important for both capitalist and socialist economies. On the one side, capitalist economies are actually mixed economies, faced with a multitude of problems of regulation and deregulation, of socialization and privatization. On the other side, many socialist economies have had mediocre success in maintaining the efficiency of their organizations, and are experimenting with the reintroduction of markets, often while trying to avoid extensive privatization. Good answers to the policy questions that face all industrialized societies depend on having empirically sound theories of the behavior of large organizations. Such theories cannot be developed from the armchair. They call for fact-gathering that will carry researchers deep into the green areas, the organizations, that dominate the terrain of our economic systems.

## Notes

* Adapted from *Journal of Economic Perspectives*, Spring 1991, 5 (2), 25–44.

1 Williamson's explanation – actually, Alfred P. Sloan's explanation (see Williamson, 1985: 10) – that employees could not be supervised adequately in their offers for used cars, is not convincing. Dealerships are also organizations, and their salesmen are employees.

2 Demsetz and Lehn (1985) cite evidence to show that corporations where ownership is widely distributed have, on average, profits as large as those with concentrated ownership. This fact does not undermine the argument of Berle and Means for conflict of economic interest; on the contrary, it raises the question – which I will undertake to answer

below – of why executives with small stakes as shareholders do appear to work for company profits

3 Everything said here about economic rewards applies equally to privately owned, nonprofit, and government-owned organizations. The opportunity for, and limits on, the use of rewards to motivate activities toward organizational goals are precisely the same in all three kinds of organizations. For sophisticated discussions of motivation and efficiency in profit-making and nonprofit organizations, see Weisbrod (1988, 1989).

4 This is not the place to describe in detail how docility and altruism induced through the docility mechanisms can be incorporated in a formal model of evolution by natural selection. I will simply sketch the general idea. Let $k$ be the average number of offspring of an individual in the absence of docility or altruistic behavior; $d > 0$ the gross increase in offspring due to docility; $c > 0$ the cost to a docile individual in offspring of the socially induced altruistic behavior; $p$ the percentage of individuals in the population which are docile and hence altruistic; and $b$ the number of offspring added to the population by an individual's altruistic behavior. Assume further, that the parentage of offspring contributed by altruism is distributed randomly through the population. Then it is easy to show that the difference between the net fitness of altruists and non-altruists (non-docile individuals), respectively, will equal $d - c$. Hence, provided that d is larger than $c$, altruists will be fitter than non-altruists. Moreover, a society will grow more rapidly the greater the fraction of altruists in it, the increase in average fitness being $(d - c + b)p$.

5 These two statements should be qualified slightly. With regard to the first, components imported by the Toyota plant from Japan may be more uniform in quality than components purchased by the other GM plants. With respect to the second, applicants interviewed for employment in the Toyota plant were screened for problem solving attitudes and skills. Note that both of these factors, whether important or not, are matters of management practice. Finally, I would not wish to claim that the factors I have mentioned were the only ones affecting the comparison.

6 Of course, perfectly competitive markets do provide stable expectations of prices at least in equilibrium, and thereby permit Pareto optima to be achieved in principle. But prices are only one of many dimensions along which uncertainty of expectations may complicate rational decision making.

## References


Berle, A.A. and Means, G.C. (1933), *The Modern Corporation and Private Property*, New York: Macmillan.

Cyert, R.M. and March, J.G. (1963), *A Behavioral Theory of the Firm*, Englewood Cliffs: Prentice-Hall.

Dearborn, D.G. and Simon, H.A. (1958), "Selective Perception: the Identifications of Executives," *Sociometry*, 21. Reprinted in H.A. Simon, *Administrative Behavior*, New York: Macmillan, 1947; 3rd edition 1976; ch.r 15.

Demsetz, H and Lehn, K. (1985), "The Structure of Corporate Ownership: Causes and Consequences," *Journal of Political Economy*, 93, 1155–77.

Devons, E. (1950), *Planning in Practice.* Cambridge: Cambridge University Press.

Hawkins, D. and Simon, H.A. (1949), "Note: Some Conditions of Macroeconomic Stability," *Econometrica*, 17, 245–8.

von Hayek, F.A. (1945), "The Use of Knowledge in Society," *American Economic Review*, 35, 519–30.

Ijiri, Y. and Simon, H.A. (1977), *Skew Distributions and the Sizes of Business Firm,* Amsterdam: North Holland.

Kornai, J. (1971), *Anti-Equilibrium*, Amsterdam: North-Holland.
Krafcik, J. and Womack, J.P. (1987), "Comparative Manufacturing Practice: Imbalances and Implications," Appendix A, working paper, International Motor Vehicle Program, MIT, May.
March, J.G. and Simon, H.A. (1958), *Organizations*, New York: Wiley.
Ross, S. (1973), "The Economic Theory of Agency: the Principal's Problem," *American Economic Review*, 63, 134–9.
Simon, H.A. (1947), *Administrative Behavior*, New York: Macmillan; 3rd edition, 1976.
—— (1951), "A Formal Theory of the Employment Relationship," *Econometrica*, 19. Reprinted in H.A. Simon, *Models of Bounded Rationality*, Vol. II, Chapter 5.2, Cambridge: MIT Press, 1982.
—— (1979), "Rational Decision Making in Business Organizations," *American Economic Review*, 69, 493–513.
—— (1983), *Reason in Human Affairs*, Stanford: Stanford University Press.
—— (1990), "A Mechanism for Social Selection and Successful Altruism," *Science*, 250.
Stiglitz, J.E. (1974), "Incentives and Risk-Sharing in Sharecropping," *Review of Economic Studies*, 41, 219–55.
Weisbrod, B.A. (1988), *The Nonprofit Economy*, Cambridge: Harvard University Press.
—— (1989), "Rewarding Performance that is Hard to Measure: The Private Nonprofit Sector," *Science*, May 5, 244.
Williamson, O.E. (1975), *Markets and Hierarchies*, New York: The Free Press.
—— (1985), *The Economic Institutions of Capitalism*, New York: The Free Press.
Winter, S. (1964), "Economic 'Natural Selection' and the Theory of the Firm," *Yale Economic Essays*, 4.

# 5 Accounting and its relationship to general equilibrium theory*

*Martin Shubik*†

Accounting both for macro and microeconomic purposes deals with process and dynamics. Much of the best microeconomic theory has dealt only with statics. General equilibrium theory shows the virtues of a price system, but abstracts from price formation and all of the accounting problems which appear in disequilibrium. An approach is suggested here for reconciliation of accounting with general equilibrium. More generally, it is suggested that the importance of accounting to economic theory has been underestimated.

The history of the development of the concepts of national income has been given by Studenski (1958) where he attributes its origin in the seventeenth century to Sir William Petty in England and Pierre le Pesant Sieur de Boisguillebert in France. He notes various theories of economic production and hence several variants of what has been called national income. In particular to this day three have survived. They are: *the comprehensive production concept* which counts goods and services; *the restricted material concept* which counts only material goods; and *the restricted market production concept* which includes only goods and services which have been delivered through the markets. Studenski suggests that Adam Smith's emphasis on the restricted material product concept may have set back the development of economics for some time.

Without delving further into the history of national income accounting it is reasonably safe to argue that by the end of World War II virtually all economists agreed upon the comprehensive production view of national income. But to this day many problems still exist in the measurement of non-market production and the gaps abound among macro-economic accounting conventions, micro-economic accounting and microeconomic theory definitions.

Far earlier than the needs of government called for the development of macroeconomic accounting, the needs of trade and production called for accounting for business. By the fifteenth century a comprehensive and logically ordered system of record keeping had become a necessity if businesses were going to grow in size and trade were to continue to expand. In 1494 the great work of Fra' Luca Pacioli (1494) laid out the basic principles of double entry book-keeping and provided the basis for the development of modem business accounting.

Both of the developments in accounting, even when regarded at the highest levels of abstraction, were and are of this world. They were directed towards helping governments and firms to cope with ongoing problems of day to day and year to year decision-making. First and foremost, accounting of all varieties has had to grapple with the dynamics of both the firm and government. A balance sheet may present a moment frozen in time, but the income statements linking two balance sheets deal directly with the dynamics of an ongoing system.

The development of macro and microeconomic theory took different routes. Macroeconomic analysis, even though partially captured by the mathematical methodology that has invaded virtually all aspects of economics, nevertheless has maintained a high content of institutional structure and a basic concern for dynamics. The causality implicit in the macroeconomic models and the predictions based upon them depend directly on a host of macroeconomic accounting assumptions. These deal with how household production is evaluated; what constitutes savings; how capital stock is evaluated; how depreciation is measured; what constitutes unemployment; how inflation is measured and many other factors usually involving both deep conceptual problems in the treatment of time and uncertainty. They also deal with fundamental economic and mathematical problems in obtaining optimal scalar or one number representations of phenomena which at best are multidimensional, or do not even have good numerical representation. An example of the measurement difficulties is provided when one tries to correct the growth of productivity for changes in the variety, quality and mixture of goods produced. If the conceptual problems, however deep, were merely confined to the academies they would not matter very much. They would increase employment for university professors and make little difference to the functioning of society as a whole. Unfortunately a cost of living index, or an inflation index or a measurement of the level of unemployment may be used to influence vital political decisions Academic economists are also citizens and as such they cannot avoid the responsibilities which must be assumed when one index is selected over another.

Many of the basic problems with the national income accounts have been covered in the paper in this journal by Richard Ruggles,[1] the emphasis here is on microeconomics. Microeconomic theory since the 1930s has developed in a considerably different manner from macroeconomics. The domain of study of microeconomics is vast and is split into many sub-disciplines. The theory of the firm, consumer choice, oligopoly theory and industrial organization maintained a high empirical content until the recent invasion of game theory, the central development in the study of the basic logic of markets, production and trade took a different turn. As the fascination with the understanding of the price system grew, progressively higher and higher levels of abstraction were considered until with the publication of Gerard Debreu's *Theory of Value* (1959) an elegant mathematical treatise presented the basic features of an efficient production and distribution of resources governed by individual optimization given the signal of an appropriate price system. The mathematics showed the conditions required for existence of such a price system but provided no indication as

to how the prices were to be attained. Furthermore it was shown that in general the prices which would clear all markets efficiently were not necessarily unique. The few verbal commentaries in the book intimated that the mathematics represented the appropriate formulation of the concepts of a competitive price system which were present in the work of many economists from the earliest times. In particular this could be taken to represent the full formalization of the writings of Walras (1954).

The development of general equilibrium theory was a great step forward in the mathematization of a central concept of economic theory. It showed the delicacy of the many conditions needed to permit the functioning of an efficient price system. It also showed the power of new and elegant mathematical applications to the social sciences. But the development of a science calls for more than mathematics and logic. The relevance of and closeness of the abstractions to the functioning economy and the nature of the questions which can and cannot be answered by the theory developed must be considered with care. General equilibrium theory is admirably suited to illustrate basic relationships in equilibrium but offers few if any insights into process. Time and uncertainty have essentially disappeared from this apotheosis of the price system. But it is time and uncertainty which are the concerns of everyday economic life and the problems of how to account for the influence of time and uncertainty in the ongoing economic process are central to the development of accounting.

Koopmans (1977) has noted that general equilibrium theory is essentially preinstitutional in character. The ability to formulate and answer in a sophisticated way many deep problems about the relationship between prices and efficient production has been of great significance in laying the groundwork for an economics that accommodates process. But this strength should not serve to hide the gap between verbal descriptions of competition and the highly noninstitutional and process free description of preferences, ownership and technological relations that make up the basic elements of general equilibrium theory.

Many of the important developments in mathematical economics in the past twenty years have been directed towards finding mathematical generalizations which weaken assumptions on preferences, take care of boundary equilibria or cast light upon production conditions A certain amount of work has been done on the stability of equilibrium points under perturbations in excess supply and demand conditions; but few general results are known concerning the dynamic of socio-economic feedback mechanisms for the formation of price.

The next step in the development of microeconomics is in building full process models showing the formation of prices via the market or showing the manipulation of prices via a centralized economic planning agency. But any attempt to do this calls for the specification of updating mechanisms of the economy regardless of any equilibrium considerations and it is precisely the recognition of the need for the description of the updating implicit in economic process that has been central to the development of accounting. Any attempt to construct full process models, even at the highest level of abstraction, forces the modeler to acknowledge the existence of, or to invent economic institutions,

because they are the carriers of economic process and in one form or the other appear as necessary consequences of fully defining "the rules of the game." The accountant takes the institutions as a fact of life. A successful post-general equilibrium theory must have them as a logical necessity of process. The growth of specific institutions undoubtedly depends on a host of factors involving history, chance, culture and technology pertaining to each society.

Ideologies often use scholarly theories as part of the social process of legitimization. The elegance, level of abstraction and apparent generality of general equilibrium theory appears to provide an intellectual justification for the extreme proponents of *laissez faire*. The market will take care of everything. Free unconstrained competition without the interference of government or society has now been "proved to be the best of all possible worlds" by advanced mathematical methods. It would be a great blessing if this were true. Unfortunately the truth lies elsewhere. The careful economic theorist with no immediate political cause to legitimize can simultaneously see the great intellectual contribution of the theory and its limitations. An understanding of the limitations may help to bring it closer to reality and in doing so may also explain why many of the definitions of the microeconomist differ from the accountant and businessman concerned with the processes of everyday production and trade.

The limitations of the general equilibrium model are as follows:

1 It is basically noninstitutional.
2 It distinguishes too few differentiated economic agents.
3 Although equilibrium conditions are specified, no explanation of price formation is given. The general equilibrium model is static.
4 There is no essential role for money.
5 The agents are modeled nonstrategically; prices are given, not formed.
6 The model is essentially timeless. Although one may make a simple *pro forma* change of notation which treats goods in different time periods as different goods, hence replaces an optimization problem in M dimensions by one in MT dimensions where M is the number of goods and T the number of the periods; this mathematical extension picks up only a few of the substantially new phenomena involved in the consideration of time and process in an economy.
7 The horizon is infinite and the agents appear without context at the start and disappear at the end.
8 The many problems posed by uncertainty in an actual economy are avoided by the same type of device that was used to extend the model to several periods. Suppose that an economy trades in M commodities and that it can be in any one of K states of nature. We may consider an economy with trade in MK commodities which may be regarded as options.
9 The proofs of the existence of an efficient price system were independent of the number of competitors, indicating that the mathematical model though appropriate for the study of efficient prices is not necessarily appropriate for the study of competitive price formation.

10 The general equilibrium model was not designed to study economies with an intermix of monopolistic, oligopolistic and competitive elements, yet every economy has this mix.
11 It is an inadequate model for the study of non-symmetric information conditions, yet disclosure of costs and other competitive information are of critical importance to the economy.
12 It has no satisfactory explanation of unemployment. In essence it offers an ideological opportunity to claim that there is no such thing as unemployment because in the best of all possible worlds without governmental or other institutional constraints the markets would adjust so that all who wished to work could find a job at the appropriate price.
13 Bankruptcy plays no role. There is neither birth nor deaths of firms. Yet a basic understanding in accounting is the need to differentiate ongoing enterprise evaluations from liquidation evaluations. Furthermore the elementary dynamics of economic process in a world with credit illustrates the proposition that the system may attain a state where prior commitments cannot be met. In such an instance bankruptcy laws become a logical necessity not merely an institutional fact.
14 Expectations are finessed. There is no learning.
15 The general equilibrium model does not need markets, banks or other credit or financial institutions, as trust is implicitly perfect. All trade is balanced at the end of the market and each individual at the start of trade has available implicitly a credit line equal to his net worth "at market," i.e., the worth of his initial assets at the final market price.
16 Firms are not institutions. Production is described by production correspondences available to all who have the resources. It is as though the only items necessary to bake a cake were the ingredients and a recipe (free to and immediately understood by all). The firm as an entity with an internal organization and a management with goals of its own is not included in this abstraction.

In the Debreu model there is really only one major actor – the consumer, and one shadow actor – the producer. The consumer has his own preference ordering and tries to maximize his own welfare. The producer is a shadowy manager of a firm which may be owned by stockholders other than himself. Even though his interests may conflict with others he is modeled as a profit maximizing perfect fiduciary who flows through the profits to a nonvoting[2] group of stockholders. The definition of profit depends upon the whole finite period being considered. There is no real concern with the accountant's need to define annual profit.

In macroeconomic models and in economic life we frequently distinguish investors, savers, speculators, brokers, bankers, consumers, manufactures, retailers and others. It is important to ask at what level of abstraction do or should these distinctions appear.

The general equilibrium model is essentially static in the sense that time is handled by merely enlarging the number of variables and renormalizing (Wold

1969). Work on sequences of economies (Green 1973) and the infinite horizon have modified this. Even so, in the work of Arrow and Debreu (1954) and Debreu (1959), no explicit mechanism for the formation of price is given, i.e., no formal markets exist. It is not surprising that at this level of theorizing one cannot distinguish a competitive economy with prices arising from competitive behavior from a socialist economy with controlled prices. The difference between them is in the mechanism of price formation; but in this theory the mechanism is left out.

In any complex economy, money plays a far more important and subtle role than merely acting as a unit of account. Yet there is no role for it in the general equilibrium model. In the past few years there have been considerable efforts made by Grandmont and Younes (1972) and many others to introduce money into a modified general equilibrium model. Although considerable progress has been made, this work is hampered by trying to stay too close to the original general equilibrium formulation.

The general equilibrium theory is nonstrategic. Prices appear by magic and the individual consumer is constrained to maximize his welfare given these prices. It is an assumption of the general equilibrium theory that the individual maximizes his welfare as though prices were given. Starting with a strategic view of competition such as illustrated by Cournot (1897), this is a *deduction* from the theory when the system is modeled as a game of strategy and solved for its noncooperative equilibria.

Points 6, 7 and 8 – that the general equilibrium model is essentially timeless, avoids the problems with exogenous uncertainty and has economic agents who are all present at the start of the economy and disappear at the end – are all essentially the artifacts of modeling the economic system.

The economic meaning to enlarging the number of different commodities traded in an economy with $M$ commodities, $T$ trading periods and $K$ states of nature, to $MKT$ is that perfect future markets and insurance contracts exist for all times and all circumstances. This is an assumption which implies perfect knowledge and trust and merits modifying.

Because the general equilibrium model is examined for the existence of an efficient price system that is independent of the number of traders, it does not offer a means for studying the effect of thin markets and few competitors. Such a distinction calls for showing how the presence of different numbers of competitors of different sizes and with different levels of control influences the formation of price. Both in law and accounting, problems involving the valuation of assets in thin or inactive markets are of considerable importance.[3]

In summary the general equilibrium system provides considerable insight into the technological and preference requirements for a society to be in a position to use a price system to achieve efficient production and distribution. As it is not concerned with process, no distinction is made between a centralized price system or one that comes about via competition among many small agents. Furthermore no insight is provided about the problems of control or the mechanisms required in either an enterprise or centralized economy. These

limitations in no way detract from the contribution to the understanding of the role of price as an efficient decentralizing device. The general equilibrium system provides a natural basis or departure point from which to consider control mechanisms. It is in the control mechanisms for economies which may have a varying intermix of competitive or controlled prices that financial instruments and institutions play their role.

The gap between the language of the economist and the accountant comes about to a great extent because even the most sophisticated general economic models are too simple to be used to analyze in any depth the myriad of *ad hoc* problems which confront the businessman and accountant in everyday life. Accounting is designed as a practical applied information processing device to aid the businessman with limited time, knowledge and data processing capability to steer his enterprise in a world laden with uncertainty. Many of its basic concepts, such as the valuation of assets, the definition of profits or the assignment of depreciation, are derived from a necessary blend of techniques of aggregation and disaggregation of information, taking into account costs of obtaining the data, together with blending in consideration of law, custom of the society and practices of the business enterprises. Like his macroeconomic accountant counterpart the business accountant is immersed in process and in assisting in every day practice. The founder of business accounting was a mathematical theorist and currently the modern professors of accounting recognize the need for a science of accounting as part of the information sciences as well as a practice of accounting, deeply reflecting the details of economic process in economies with different legal systems and customs. The paper of Yuji Ijiri in this journal[4] serves to illustrate that the levels of abstraction and concerns in the development of accounting theory are close to those in microeconomic theory.

The gap between modern accounting theory and practice and economic theory is large and to a great extent exists because the need to reconcile the two was less important in simpler economies. The reconciliation must come from both ends, like the digging of a major tunnel. In this note only the economists' responsibility is stressed.

Much of this note has been devoted to a critique of general equilibrium theory, yet paradoxically this is because of my belief as to how important and valuable it is to the task noted above. I suggest that general equilibrium theory produced a considerable insight into the nature of some of the basic features of the price system. The rigorous standards of logical consistency and completeness in the characterization of an extremely complex problem were obtained at the cost of considerable simplification. In particular all traces of economic process were eliminated from the basic structure studied. But the next step is to transform and extend the models of general equilibrium theory into process models. This can be done utilizing game theory following the work of Cournot (1897) to its more modern manifestations (Shubik 1973; Shapley and Shubik 1977; Dubey and Shubik 1978). The strategic market game is a microeconomic process model of the economy. Both the modeling and the analysis force the construction of a *mathematical institutional economics* where apparently minor institutional

detail such as bankruptcy and default rules or details concerning accrual accounting turn out to be logical necessities in being able to well define process. By insisting on process models which are well defined playable games, the modeling emphasis is away from equilibrium and calls for the description of carriers of process which in turn can be interpreted as elementary market and financial institutions. In such a formulation virtually every concern of the practicing accountant and financial analyst, be it seniority of claims, definition of periodic profits, treatment of inventories, evaluation of "haircuts" for risky loans or treatment of strategic opportunities as off-balance sheet assets or liabilities, appear as needed rules to facilitate process.

From a purely practical point of view the reconciliation between economic theory and both macro and micro accounting may come about to a great extent by the growth in sophistication of the economic agents who understand that there are great fortunes to be made in arbitrages between accounting and legal conventions which do not match the underlying economic realities sufficiently closely. Thus in the early development of commodity and stock markets the brokers could earn considerable incomes off the escrow accounts or delays in settlement of the balances of their clients, and to this day what constitutes a legal transaction in one jurisdiction may be interpreted otherwise in a different jurisdiction. Further examples are provided by an old New York observation that the wise businessman does not maximize wealth, but undisclosed wealth, and by noting that an accounting convention which treats an investment as an immediate expense may completely transform a tax bill.

Keynes (1936) observed many years ago that: "Practical men, who believe themselves to be quite exempt from any intellectual influences, are usually the slaves of some defunct economist." It is undoubtedly true that the practicing businessman and accountant has little use for the economic theorist, but the theorist, if he wishes to live up to the praise of Keynes, must understand the world as it is. He must be able to fit Generally Accepted Accounting Principles and their European and other equivalent codes into his theorizing. He must remember that the rules of thumb, approximations and customary activities of the business community are not merely irrelevant ephemera to the theorist, but their peculiarities and paradoxes must at least be shown to be special cases of a general theory capable of explanation by the theory. If this is not the case, then the theory not the reality is at fault.

## Notes

* Adapted from *Economic Notes* by Monte dei Paschi di Siena, 22 (2), 1993, 226–34.
† Department of Economics, Yale University, New Haven, CT, USA.
1 R. Ruggles (1993), National Income Accounting: Concepts and Measurement. Economic Theory and Practice, *Economic Notes* by Monte dei Paschi di Siena, 22 (2), 1993, 235–64.
2 It is hard to give much meaning to stockholder voting in the general equilibrium model without specifying a great amount of corporate law and running the risk of losing price system if the laws are not appropriate (Shubik's note).

3 Pantaleoni (1904): *Alcune osservazioni sulle attribuzioni di valori in assenza di formazione di prezzi di mercato* [Some observations on the imputations of values when price formation lacks], *Giornale degli Economisti*, March, 203–31, and April, 307–25, early recognised.
4 Y. Ijiri, "The Beauty of Double-Entry Bookkeeping and its impact on the Nature of Accounting Information," *Economic Notes* by Monte dei Paschi di Siena, 22 (2), 1993, 265–85.

## References

Arrow, K.J. and Debreu, G. (1954), "Existence of an Equilibrium for a Competitive Economy," *Econometrica* 22, 265–90.

Cournot, A.A. (1897), *Researches into the Mathematical Principles of the Theory of Wealth,* New York: Macmillan (Bacon's translation of French 1838 original).

Debreu, G. (1959), *The Theory of Value,* New York: Wiley.

Dubey, P. and Shubik, M. (1978), "The Noncooperative Equilibria of a Closed Trading Economy with Market Supply and Bidding Strategies," *Journal of Economic Theory,* 17 (1), 1–20.

Grandmont, L.M. and Younes, Y. (1972), "On the Role of Money and the Existence of a Monetary Equilibrium," *Review of Economic Studies,* 39.

Green, L.R. (1973), "Temporary General Equilibrium in a Sequential Trading Model with Spot and Futures Transactions," *Econometrica,* 41.

Hahn, F. (1973), "On the Foundations of Monetary Theory," *Essays in Modem Economics*, (Michael Parkin and A R. Nowbay, eds), New York: Harper and Row.

Keynes, J.M. (1936), *The General Theory of Employment, Interest and Money,* London: Macmillan.

Koopmans, T.C. (1977), "Concepts of Optimality and their Uses," *The American Economic Review,* 67 (3), 261–74.

Pacioli, L. (1494), *Summa de Arithmetica, Geometria, Proportioni et Proportionalita.*

Shapley, L.S. and Shubik, M. (1977), "Trade Using one Commodity as a Means of Payment," *Journal of Political Economy,* 85 (5), 937–68.

Shubik, M. (1973), "Commodity Money, Oligopoly, Credit and Bankruptcy in a General Equilibrium Model," *Western Economic Journal,* II, 1, 24–38.

Starr, R. (1974), "The Price of Money in a Pure Exchange Economy with Taxation," *Econometrica*, 42, 45–54.

Studenski, P. (1958), *The Income of Nations*, New York: New York University Press.

Walras, L. (1954), *Elements of Pure Economics*, Homewood: R.D. Irwin (W. Jaffe's translation of the 1874–1877 original, *Élements d'économie politique pure*, Lausanne: L. Corbaz).

Wold, H.O. (1969), "Mergers of Economics and Philosophy of Science," *Synthese*, 2, 427–82.

# 6 Accounting and the theory of the firm*

*Ronald H. Coase*†

This paper describes the background and objectives of a series of papers written fifty years ago at the London School of Economics (LSE). One objective was to encourage the use of accounting numbers in economic research. A second objective was to improve the theory and practice of accounting. Understanding cost accounting and opportunity costs within a firm was allied to understanding the organization of firms. The theory of the accounting system is part of the theory of the firm. Like a similar request made fifty years ago, the paper concludes with a call for inter-disciplinary studies between economics and accounting.

This paper will be largely autobiographical and will mainly be centered around what happened at the LSE fifty or more years ago. In 1937 my article "The Nature of the Firm" was published, based on ideas I had developed some years earlier. It has since been reprinted on several occasions and has been much discussed, particularly in recent years. One year later, in 1938, I wrote a series of articles for *The Accountant* entitled "Business Organization and the Accountant," articles which have also been reprinted and referred to in the literature,[1] although not nearly to the same extent as "The Nature or the Firm." At any rate, in the 1930s I was working on both accounting and the theory of the firm. Today I want to discuss the relation between these subjects.

I explained how I came to be interested in and to write on the theory of the firm in some lectures on "The Nature of the Firm" which I gave at Yale last year (Coase 1988) and which some of you may know. I will now explain how I came to be interested in and to write on accounting. In my Yale lectures I disclosed, which may be somewhat surprising given that my career has been that of an economist, that I did not take any course in economics while I was a student at the LSE, a circumstance which I believe gave me a freedom in thinking about economic problems which I might not otherwise have had. The position with regard to accounting is quite different. I took a degree in commerce, with the thought that I would go into business. As I chose the "Industry" group in my final year, an option supposedly designed for those who were going into works management, I took courses in works and factory accounting and cost accounting as well as financial accounting. I learned, among other things, that there

were many methods of calculating depreciation, valuing inventories, allocating on-costs, and so on, all of which gave different results but all of which were perfectly acceptable accounting practices. In addition, there were, of course, practices which were not acceptable and which should not be countenanced by any accounting firm which aimed at respectability.

As we were to discover, the views that I and my young colleagues at the LSE came to hold were considered by a number of accountants (including the head of the accounting department at LSE) not to be respectable and, indeed, it would not be going too far to say that they were thought to be sacrilegious. The books that had been entrusted to the accountants' keeping were apparently sacred books. Those most involved with this critique of accounting practices were Ronald Edwards and Ronald Fowler. Edwards, after the war, became Chairman of the Electricity Council, which was responsible for the operations of the British electricity supply industry, and later became head of Beechams, the pharmaceutical company. But in the middle 1930s he was a humble assistant lecturer like myself. He had worked in an accounting firm and taken the external B.Com., studying with the Commerce Degree Bureau (which organized correspondence courses under University auspices). His work as a student with the Commerce Degree Bureau brought him to the attention of Arnold Plant and, as a result, he was appointed in 1935 to a position in the newly formed Department of Business Administration at LSE.

Ronald Fowler, who had been a fellow student of mine, wrote a book in 1934 on *The Depreciation of Capital* and worked with Plant on a project which surveyed annually the costs of British department stores (an investigation modelled on a similar one carried out by the Harvard Business School). Fowler joined the Department of Business Administration after Plant became its head. As a result of his work on depreciation, Fowler developed a keen interest in accounting theory.

Edwards played the leading role in our study of accounting and the work of accountants, and following his initiative the Accounting Research Association was formed in 1936, with the active cooperation of the Librarian of the Institute of Chartered Accountants, Cosmo Gordon. The work with which I was associated had two aims. The first was to show economists that the accounts provided valuable source material for economic research, and the other was to persuade accountants to change their practices so as to make the accounts more valuable for this purpose by, for example, not lumping together disparate items and by adopting more uniform methods which would facilitate the aggregation of the accounts of different companies. We were not content merely to stop there but went on to show economists what could be done and accountants what should be done. We did this by making an economic study of the British iron and steel industry, using the accounts of public companies.

The paper "Published Balance Sheets as an Aid to Economic Investigation – Some Difficulties" came out in 1938 and recounted the difficulties we had encountered in carrying out our investigation. We started by asserting that "balance sheets, in so far as they are reliable and the conventions on which they

are based are thoroughly understood, will provide the economist with much material for applied study" and we pointed out that what we said was based on an examination of approximately 2,000 balance sheets. In summing up, we said that in drawing attention to the difficulties in interpreting the various items in the balance sheet, we were not claiming that "the economist could do better than skilled accountants. The aim of this paper is merely to emphasize the problems as they are seen to users of balance sheets for economic research." We concluded that "some of the major difficulties are due to the lack of uniformity in the practice of accountants, and also, which is to some extent the reason for this, the obscurity of much of their reasoning and the distinctions which they employ. It follows that an improvement in accounting theory would materially assist the work of economic research."

Later in 1938 we presented our findings on the iron and steel industry to the annual meeting of the Association of University Teachers of Economics. It aroused little interest. This was no doubt in part due to the Chairman's introduction. Usually the Chairman's introductory remarks are so flattering that one's heart sinks as one hears them, realizing that now whatever one says must inevitably be a disappointment. In our case, the Chairman's introductory remarks had a different character. They were uniformly hostile. I cannot now remember what they were, but they expressed the view that the figures provided by accountants were so unreliable that any attempt to use them for economic research was vain.

Nevertheless, we continued with our work and in late 1939 or early 1940 "The Iron and Steel Industry 1926–35: An Investigation Based on the Accounts of Public Companies" (Special Memorandum No. 49 of the London and Cambridge Economic Service) was published. Of course our methods of analysis were extremely crude compared with those employed today. And the fact that there were no computers severely limited the calculations that could be made. All our computations were made by pencil and paper, with the aid of a Monroe hand-calculator. Nonetheless, whatever the value of our results, I think we showed the potential usefulness of these accounting figures for economic research. I will give as an example of what could be done the discussion of the factors determining the demand for bank advances (one of the parts of the study for which I was responsible).

It was then widely believed that bank credit played a crucial role in generating and intensifying the business cycle and that this was particularly the case in the capital goods industries, industries which, in the view of some economists, played a dominant part in the whole process. The iron and steel industry can be regarded as the archetype of the capital goods industries. Our analysis of these balance sheets showed that, at least so far as our sample was concerned, the beliefs of some prominent economic theorists about the behaviour of bank credit were incorrect. Our figures confirmed the common view that most borrowing from British banks was for relatively short periods, but in other respects the results were quite unexpected. Analysis of the balance sheets showed that increasing bank credit was associated with losses and declining profits. The

result was that, although with the onset of the depression in 1929 bank credit fell immediately, as the depression deepened bank advances to the iron and steel industry rose; and when, in 1935 and 1936, profits began to increase significantly, bank credit fell. Bank credit in the iron and steel industry, far from exaggerating the effects of a business fluctuation, acted as a cushion and as a moderating force.

In drawing attention to this part of our work, I do not wish to imply that this was our most significant finding. I have mentioned it because it was a part of the investigation for which I was largely responsible and with which therefore I have greater familiarity. Other parts of the investigation dealt with the movements of profits, undistributed profits, liquid balances, gross working capital, stocks, work in progress, and so on, and some of these sections are no doubt more important than that on bank credit.

We felt able to conclude that "published accounts are a source of economic statistics that ought not to be neglected." This view resulted in large part from the fact that, by using the accounts to derive information about firm behaviour, it was possible to relate the figures about one kind of firm behaviour to all other aspects of the firm's position and behaviour. This still seems to me a decisive advantage of using the accounting records.

The war intervened in 1939, and when it ended we did not return to this work. Whether similar work has since been carried out by others I do not know, but if it has, so far it seems to me that it has little effect on economic theory. Certainly the vastly improved quality of published accounts (many of the difficulties mentioned in our Accounting Research Association paper would not now exist), as well as the improved methods of analysis, should make it possible today to use the published accounts more fruitfully than we were able to do.

I would not wish to leave you with the impression that the Accounting Research Association was mainly concerned with the use of accounting records for economic research. It was also, and probably more, concerned with the improvement of the theory and practice of accounting. As part of this effort, it conducted, by means of questionnaires, studies of the methods of valuing raw materials, work in progress, and finished stock and of the methods of calculating depreciation actually used in business – which enabled an attempt at discovering how the method employed varied with the type of business.

In the examination of the financial accounts, I was very much the junior partner. Edwards knew about actual accounting practices, of which I was completely ignorant. And both Edwards and Fowler knew more about accounting theory than I did. My name appears first simply because in the alphabet C comes before E and F.

I now turn from our work on financial accounting to that on cost accounting, and in our discussions of this subject I played a more equal role. My published contribution consisted of twelve articles which appeared weekly in *The Accountant* in 1938, entitled "Business Organization and the Accountant." These articles were written each Wednesday night and were published on Saturday. They could be written in this way because I thought of them not as presenting ideas

of my own but as an exposition of views which were the common property of the economists at LSE or, at any rate, of those with whom I was associated.

That the ideas in these articles were in no sense original is apparent if one also studies Edwards' "The Rationale of Cost Accounting," which appeared in 1937 in *Some Modern Business Problems* and which has been reprinted elsewhere, or the paper of C.L. Paine (another student of Plant's), "Some Suggestions for Measuring the Relative Profitability of Competing Uses of Floor Space and Fixed Equipment," or the writings of David Solomons, such as his "Cost Accounting and the Use of Space and Equipment" (published in 1948 and reprinted in *Studies in Costing*). Solomons, of course, studied at LSE under Arnold Plant, and in a letter which Fowler wrote to me in 1932 while I was in the United States and which I have deposited in the Regenstein Library of the University of Chicago, he refers to a very good paper on costs which had just been given in Plant's seminar by a student, David Solomons.

Originally, I had not intended to write a series of articles in *The Accountant* dealing with costs and cost accounting. Edwards had suggested that I write a series dealing with problems of business administration. But when, in the summer of 1938, I sent him a draft of an article on vertical integration (whether to make or buy), Edwards complained bitterly that the accountants for whose benefit the articles were supposedly written would not understand what I was talking about, as the concepts and terminology I used would be completely foreign to them. I therefore decided to write an introductory section in which I explained the character of my approach. However, the introductory section came to occupy the whole of the twelve articles, and business problems were not discussed except as illustrations of the value of my approach.

When I wrote these articles I had no idea that they would ever be reprinted, as they were by Solomons in *Studies in Costing* and in *Studies in Cost Analysis* and by Thirlby and Buchanan in *LSE Essays on Cost*, nor that they would ever be referred to in the literature. Had I known this, I would undoubtedly have tried to be more careful in the writing – with the probable result that these articles would never have been completed.

What these extremely simple articles did was to argue that in business administration cost should be interpreted as opportunity cost and that the approach should be marginal (concerned, that is, with variations in costs and receipts). This led naturally to a critique of the practices and doctrines of cost accountants. The opportunity cost concept developed at LSE was of course derived from Knight and Wicksteed as expounded by Lionel Robbins, and was also no doubt influenced by Hayek who would have added an Austrian flavour.

As I have indicated, accounting as taught at LSE at that time had the aspect of a religion. It is not, therefore, altogether surprising that these articles resulted in my being given a public rebuke by S.W. Rowland, the head of the accounting department at LSE, in an address given to the Accounting Research Association in 1939, entitled "Experience, Research and Speculation in Accounting." He opened his reference to my articles with the following sentences: "lf I particularize I do so with pure objectivity. I notice cases where the impact of economic

studies on accounting lays greater emphasis on the speculative element than on experience." He then proceeded to ridicule the opportunity cost concept after completely misunderstanding what I had said. On other occasions, in *The Accountant* he attacked in a vitriolic way the views of Edwards on the nature and measurement of income. In an article he compared Edwards' ideas to Dr. Johnson's leg of mutton, "ill bred, ill fed; ill killed, and ill drest," and in a letter in the correspondence columns he stated that he regarded these ideas "as dangerous nonsense made the more dangerous by the fog of words in which assumptions are disguised as truths."

Later, in 1947, W.T. Baxter, who had written two papers on "The Form of Final Accounts" for the Accounting Research Association, was appointed head of the accounting department at LSE and relations with the economists became very cordial. I should add that my articles also brought some criticism from accountants in the correspondence columns of *The Accountant*. However, we also found that we had some allies among cost accountants, and I remember particularly those at the Gramophone Company (which was or became part of EMI). Of course, we would tend to hear from those cost accountants who supported our efforts.

Buchanan, in his book *Cost and Choice*, has argued that the concept of cost developed at LSE in the 1930s was special to that institution, and he claims that my articles embody this particular point of view. He instances with approval my linking of costs with the decision to be taken. I "tie[d] cost to choice," as Buchanan puts it. He also notes my rejection of any classification of costs into categories such as fixed and variable "independent of the identification of the decision under consideration," and he makes other points of a similar character. Whether Buchanan is right in his thesis that these views of cost were not those normally held by economists elsewhere I do not know. But there is no question that I thought of these articles as presenting views which I shared with my economist colleagues at LSE. As I said in an introductory note when these articles were reprinted in LSE Essays on Cost: "If Professor Buchanan's thesis about the special character of the LSE approach to costs is correct, it is the fact that these articles do not represent a personal point of view which gives them their historical significance."

The opportunity cost concept of the London tradition of the 1930s which Buchanan had praised was to come under attack from a somewhat unexpected quarter. In 1974, a collection of papers entitled *Debits, Credits, Finance and Profits* was presented to W.T. Baxter on his retirement from the chair of accounting at LSE. J.R. Gould, a lecturer in economics at LSE, contributed a paper in which he discussed Buchanan's thesis and the London tradition. Although his arguments were essentially the same as those he had put forward in 1962 in *Studies in Accounting Theory*, in 1974, no doubt stimulated by Buchanan's book, he made detailed references to the writings of Edwards, Baxter, Solomons, and myself. He says this of my use of the opportunity cost concept: "Coase's specification of opportunity cost, if interpreted as a definition for use in computation, is at best superfluous, and at worst downright misleading." It is not a statement that one would be inclined to choose for one's epitaph.

Gould gives, as an example of how my definition could mislead, the case of a "somewhat obtuse developer" who interprets my definition of opportunity cost in a business decision as the biggest alternative receipts to mean the biggest gross receipts and who therefore does not deduct the costs that would have to be incurred to secure those receipts. This does not seem to me a substantial objection to the opportunity cost concept nor to its use in computation. Gould continues: "A less obtuse developer would interpret opportunity cost to mean highest alternative net receipts. ... The correct decision would be made – but the concept of opportunity cost has played no useful part in the computations necessary to arrive at the decision. The net receipts of each plan must be worked out before the opportunity cost of any one plan can be computed; and once the net receipts are known the correct decision can be made without further computation."

This is an interesting point, but consideration of it leads, I believe, to a conclusion different from that drawn by Gould. For one thing, if net receipts, that is, receipts minus costs, are to be calculated correctly, costs must be calculated correctly, that is, they should represent opportunity costs. If, for example, one of the plans requires the use of a material already in stock in that business and that material cannot be sold and is of no use elsewhere in the business, it would obviously be wrong to treat the cost of using that material as being equal to the price that was paid for it in the past. It is, in fact, zero.

All this I explained in "Business Organization and the Accountant." But there is a more serious objection to Gould's way of looking at the problem, and it is one which takes us to the heart of the theory of the firm.

When I was starting to study economics some 60 years ago, it was quite usual to illustrate the economic problem by considering how Robinson Crusoe would make his decisions. He had all the resources of his island at his disposal, but they were limited and he would not be able to achieve all he would like. He had to choose. He would choose to utilize these resources in those combinations which would yield that collection of products out of all possible collections that, in total, he preferred. Some modern economists would say that he maximized his utility. In reaching this position, Robinson Crusoe need make no explicit calculation of cost, although he might find it useful to place on his bulletin board "one beaver equals two deer" or perhaps in his case "fifty coconuts equals one turtle."

Today, one would not be likely to start with this isolated man but with a society containing many men. The economic problem faced is then not simply to decide how one individual would use resources but also which individual should use them. It is normally assumed that this society uses money and that the actions of its members in the economic sphere are coordinated by a pricing system. To secure the use of a resource it is necessary to pay at least as much as someone else would. Consequently, prices are bid up until the amount demanded equals the supply. A consumer faced with these prices would make his purchases of the various goods and services so as to maximize utility. An individual who used resources to make a product for sale would aim to maximize

profits, the difference between what he would receive for his product and what he had to give up to obtain the resources required to make it. Costs become explicit as a result of the existence of the demand of others for resources. However, the prices of the products he sells and of the resources he uses are not the result of any calculations he makes but are given to him as a result of the operations of the market. Since, leaving aside the effects of monopoly, the prices paid for resources must be equal or (slightly) greater than they would yield in another use or to another user, cost (the price of the resources) is opportunity cost, and resources will be employed in such a way as to maximize the value of production.

This is probably the way that most modern economists look at the working of the economic system, and it undoubtedly brings out some of its most important features. But, as I've often said, it assumes that the operation of the market is costless, which it is not. Resources are not employed in those uses where they make the most valuable contribution, not because of the effect of such factors as the cost of transport, which is commonly taken into account, but because of the existence of what has come to be called transaction costs. I have argued, and I need not repeat the argument here, that firms emerge because they are able to achieve some of those more productive arrangements of factors of production which, because of their cost, are made impossible if one relies solely on market transactions.

I said in "The Nature of the Firm" (and I have not changed my mind) that the expansion of a firm will halt at the point at which the costs which it has to incur to organize an additional transaction within the firm become equal to the costs of carrying out that same transaction on the market or to the costs of organizing it within some other firm. But what determines where this point will be? It cannot be said that we see very clearly the answer to this question, although as a result of the work now going forward the mists are beginning to clear. My present feeling is that, while transaction cost considerations undoubtedly explain why firms come into existence, once most production is carried out within firms and most transactions are firm–firm transactions and not factor–factor transactions, the level of transaction costs will be greatly reduced and the dominant factor determining the institutional structure of production will in general no longer be transaction costs but the relative costs of different firms in organizing particular activities. This does not mean that transaction costs will not be important in particular cases nor that they will not be important in determining the form of the contractual arrangements made by firms. What it does mean, if I am right, is that, as I put it in my Yale lectures, "to explain the institutional structure of production in the system as a whole it is necessary to uncover the reasons why the cost of organizing particular activities differs among firms."

This leads us to a point of particular interest to the participants at this conference. If economists are to study the determinants of the costs of organizing various activities within firms, they will have to call in the assistance of accountants since the costs of organizing clearly depend on the efficiency of the

accounting system. It is in this connection that Gould's view that the opportunity cost concept is not useful for business computations, since all that is required is to maximize profits, seems to me wrong. In a firm, men also make decisions independently of those that others are making (there is delegation of responsibility and all decisions are not the product of a single mind) and when people in a firm use its resources, they often need to be given some figure representing their costs, so as to be able to compare it with receipts. Since using a resource denies its use to others, the figure for cost should represent what it would yield elsewhere in the business. In this planned society, the firm, costs do not, in the main, arise directly out of the operations of the market but are computed and provided by the accounting system. While outside the firm prices and therefore costs are explicit (because of the demands of others for resources) and are determined by the operations of the market, within the firm there are explicit costs for exactly the same reason but they are provided by the accounting system. This internal system takes the place of the pricing system of the market.

In "Business Organization and the Accountant" I was concerned that these accounting calculations be properly made. But doing this is, no doubt, more difficult in some circumstances than in others – and we need to know what these circumstances are. When it is difficult, because of the particular activities or combination of activities in which the firm is engaged, the costs of organizing will be greater – either more mistakes will be made or additional costs will be incurred to avoid making them. As a result, the activities which we find firms undertaking must be influenced to some degree by their effect on the efficiency with which the accounting system operates.

Most production in a Western economy is carried out in this planned society, the firm, and when a firm decides not to take on another activity, what normally happens is that it will be done by another firm, if it is done at all. In a footnote in "The Nature of the Firm," which so far as I know has never been commented on, I argue that in a competitive system there is an optimum amount of planning, and, I could also have added, an optimum quantity of market operations. In understanding how in a competitive society the choice is made between these alternative but interrelated means of organization, we must take into account the role of the accounting system. The theory of the accounting system is part of the theory of the firm. It is not my belief that the secret to the determination of the institutional structure of production will alone be found in the accounting system, but it certainly contains part of the secret.

I suggested that the accounts could be a valuable source of data on firm behaviour, and, if I am right, it follows that their use could greatly assist in the development of a theory of the firm. Then I argued that a theory of the accounting system is part of the theory of the firm. If this view comes to be generally recognized, we may expect to see a growth in interdisciplinary studies between economics and accounting. I hope this happens, and it would certainly give me great pleasure to take up once again the investigation of those interrelations between economics and accounting about which I and my colleagues at LSE held such high hopes some fifty or more years ago concerning what

could be achieved. What I have foremost in my own mind is the gain to economics that would result. But it seems reasonable to suppose that an increased understanding of the part which accounting plays within the firm will not be without some advantage to accounting.

## Notes

* Adapted from *Journal of Accounting and Economics*, 12 (1990), 3–13 (received December 1988, final version received February 1989).

† University of Chicago Law School, Chicago, IL, USA.

1 Coase, H. Ronald (1981), shortened version of a series of twelve articles from *The Accountant*, 1.10.1938 – 17.12.1938, reprinted in Thirlby and Buchanan (1973).

## References

Baxter, W.T. (1939a), "The form of final Accounts: Part A," Accounting Research Association publication no. 4.

—— (1939b), "The profit and loss account," Accounting Research Association publication no. 4.

—— (1940), "The form of final accounts: Part B," Accounting Research Association publication no. 8.

Baxter, W.T. and Davidson, Sidney (1962), *Studies in accounting theory* (Sweet and Maxwell).

Buchanan, J.M. (1969), *Cost and choice* (Markham).

Coase, Ronald H. (1937), "The Nature of the Firm," *Economica* 9, 386–405.

—— (1988), "The nature of the firm: Origin, meaning, influence," *Journal of Law, Economics and Organization* 4, Spring, 3–47. (Reprinted in O.E. Williamson and S.G. Winter, eds, *The Nature of the Firm. Origins, Evolution, and Development*, Oxford: Oxford University Press, 1991.)

Coase, R.H., Edwards, R.S. and Fowler, R.F. (1938), "Published balanced sheets as an aid to economic investigation: Some difficulties," Accounting Research Association publication no. 3.

—— (1939), "The iron and steel industry 1926–35: An investigation based on the accounts of public companies," London and Cambridge Economic Service special memorandum no. 49.

Edey, Harold and Yamey, B.S. (1974), *Debits, credits and finance and profits* (Sweet and Maxwell).

Fowler, R.F. (1934), *The depreciation of capital, analytically considered* (P.S. King and Son).

Paine, C.L. (1938), "Some suggestions for measuring the relative profitability of competing uses of floor space and fixed equipment," Accounting Research Association publication no. 2.

Plant, Arnold (1937), *Some modern business problems* (Longmans, Green and Co.).

Rowland, S.W. (1938), "The nature and measurement of income: A rejoinder," *The Accountant*, 483.

—— (1939), "Experience, research and speculation in accounting," *The Accountant*, 442–4.

Solomons, David (1948), "Cost accounting and the use of space and equipment," *The Accountant*, 240–1.

Thirlby, G.F. and Buchanan, J.M. (1973), *LSE essays on cost* (Weidenfeld and Nicolson), 95–132. New York and London: New York University Press, 1981.

# 7 The impact of the corporation on classical economic theory*†

*Adolf A. Berle Jr*

Some measurable fact-phenomena: size and scope of large corporate activities; distribution of ownership; change in wealth-holding; source of and power over capital. The shift from "capitalist" control. The immutability of classical economic principles: competition; maximization of profit. Stockholders derive what influence they have from social-political, not from entrepreneurial, factors.

## Introduction

In 1932 the thesis was presented by myself and Gardiner C. Means that the growth and functioning of large corporations introduced certain elements not adequately taken into account by classical economic theory.

One such element was the shift of management function away from entrepreneurial "capitalist" owners and to administrators; another, that there was in process an inevitable alteration in the position of shareholders, changing the traditional logic of property as respects "ownership" of these corporations. As these trends continued, the shift would have increasingly greater effect.

Continued observation thereafter indicated increasing intrusion of at least three other developments: (a) the competitive process was changing in quality, impact and effect; desire for market-power increasingly was becoming the controlling consideration; (b) formation and control over application of capital was increasingly ceasing to be individual and (where not carried on by the state) was increasingly becoming a function of the large corporations. Capital formation itself was increasingly effected by corporations through price rather than through personal savings, in view of corporate capacity to include as part of its price not only depreciation allowances but also an item of profit not designed for distribution to stockholders; and (c) finally, the role of the stock and securities markets as sources and allocators of capital was declining, notably in the case of risk capital and markets for equity stocks.

As corollary, large corporations increasingly would come to be regarded, and to regard themselves, as part of a political-economic system rather than as classical merchant adventurers.

Properly, these propositions are now reviewed by economists. So, of course, they should be. Particularly, the neoclassical school of economic thought rejects the idea that any change in theory is required by current phenomena.

To this writer, the neoclassically oriented critiques of the propositions mentioned above seem not to take adequate account of the factual results of the flood-tide of institutional development which carries the bulk of the burgeoning industrial evolution. Scant heed is paid to the vast (the word is used advisedly, not rhetorically) changes in productive and commercial processes. Too little attention is paid to the changes in quantity, quality, content and distribution of the resulting ownership interests. In neoclassical theoretical analysis, there has been a natural, though unhappy, tendency to use classical economic terms and phrases (accurate when used to describe conditions of half a century ago) as though their then-content accurately describes today's processes.

No one denies that the bases of the present system are "capitalist" in origin. But to assume from that historical fact that "capitalism" is the same system as that prevailing, let us say in 1900, is about as relevant as to assume that a modern motorcar is essentially the same as a fringed surrey because both have four wheels and transport passengers. Still less is it sound to conclude (as does Professor Peterson)[1] that merely because the American economy is mainly dependent on "voluntary, self-supporting private enterprise," that fact "largely precludes serious departure from the other principal features of capitalism as traditionally viewed." That proposition involves an attempt to maintain that "capitalism," as classically understood, has not evolved to the point of change, despite the huge volume of factual, technical and statistical evidence to the contrary.

The writer does not hold himself responsible (despite Peterson's inclusion of them) for all projections, deductions, speculations which have been drawn from the phenomena of corporate development by a growing number of observers and commentators, though all are interesting, some are important, and few can be safely ignored. This essay deals merely with the salient points of Peterson's neoclassical thesis. These appear to be:

1 No significant alteration has occurred in the location of managerial responsibility or in the ownership-control thereof, requiring change either in economic theory or its application.
2 In any case, the fundamental of capitalism remains unchanged: the motivation and practice of corporations remains that of profit-maximization, and they remain controlled by competition and cognate market forces so that significant intrusion of social motive is and must be of negligible effect.
3 No significant change has occurred in the institution of "property" as represented by stock held by stockholders, and the stock markets in substantial measure retain their capital allocation function.

His conclusion appears to be that investment, production and distribution, and the position of ownership are all motivated, carried on and maintained in the traditional way. For neoclassical economists, business remains as usual. Given the facts, I think they are wrong.

## Preliminary: Some measurable fact-phenomena

As preliminary, it is not inapposite to call attention to a few statistically measurable phenomena in all three fields. If stereotypes of economics, musty or otherwise, are being demolished, their destruction has been accomplished by observable fact rather than by "wayward" commentators.

### *Size and scope of large corporate activities*[2]

For the year 1963, the 500 largest industrial corporations had combined sales of $245 billions; these accounted for about 62 per cent of all manufacturing sales in the United States. (The largest of these 500, General Motors, accounted for $16,500,000,000; the smallest, $86 millions.) Surrounding most of the giants is a penumbra of nominally independent but actually captive, or dependent, or market-controlled companies whose market decisions and behavior move more or less along lines determined by the central large-scale corporation. This multiplying factor does not show up in the figure given. Few fair-minded scholars would deny that big corporations dominate the manufacturing scene. Obviously, in greater concentration, large corporations even more markedly dominate the transportation, public utility and communications industries.

In all the 500 largest corporations (there are a handful of atypical exceptions such as Du Pont, Ford, Time, Inc., and some smaller oil companies) and a number of the smaller corporations (whose market percentage is relatively statistically small), "control" is atomized among large stockholder lists ranging from a minimum of several thousand to a maximum of more than a million. This process of atomization is not complete, but it is continuously going forward.

### *Distribution of ownership*[3]

The total number of individual stockholders is estimated at between 17 and 20 million individuals (more probably the lower figure). In addition, financial institutions (pension trust funds, mutual investment funds, fire insurance companies and others) have substantial holdings of stock; these are held for a far larger number of individuals who derive income or other benefits through these institutional conduits. An extremely rough estimate of the number of these individuals would add not less than 25 million more to the figure (this is a drastic underestimate) though they appear as pension trust beneficiaries, etc., and not as individual stockholders.

### *Change in wealth-holding*[4]

The over-all change in the property system forecast more than thirty years ago has gone far towards reality. At the close of 1963, total personally-owned wealth in the United States was estimated by the economic department of the First National City Bank (on the basis of Federal Reserve and National Bureau of

Economic Research statistics) at $1,800 billions. The largest item, $550 billion, consisted of corporate stock. The other items were: $200 billion of life insurance company reserves, United States securities, corporate bonds and the like; and $375 billion of liquid assets such as cash and bank deposits. These three items total about $1,125 billion – just under two-thirds of the total personally-owned wealth. (The balance consisted of owner-occupied homes and personally-owned durables.) *Nearly one-third of all personally-owned property, apparently, now consists of stock, representing ownership of the corporate system.* More, indeed, if the individually-owned indirect holdings of stocks, bonds and securities, held chiefly through pension rights and fiduciary institutions, are included. Further, if anyone cares to follow the statistics over the past two decades, it will be apparent not only that personally-owned wealth has absolutely increased, but also that the elements of its make-up have undergone a major change as stock increasingly replaces personally-owned "things." The word "revolutionary" in its current, rather weakened, sense is not inapplicable to that change.

### *Source of and power over capital*

Finally, though Peterson pays disproportionately little heed to this, a more striking shift has occurred in the method of accumulating and the decisions governing the application of, capital. More than 60 per cent (probably converging this year toward 65 per cent) of all capital entering industrial corporations is internally generated by accumulating depreciation allowances and undistributed profits – both items being produced by charging to the consumer prices sufficient to permit such accumulations. Another 20 per cent of such capital is derived from bank credit extended directly by commercial banks or industrial corporations, presumably in anticipation of such accumulation. Only the balance – not more than 20 per cent at best-and probably closer this year to 15 per cent – is derived from personal "savings." These conceivably might be material for classical risk-taking, decision-making or other capitalist application by their owners. Factually, they are not. They go overwhelmingly into intermediate institutions such as savings banks which perform this function, and are not applied to risk-capital investment. So much of this item as goes into equity or risk capital operations is largely devoted to a single group of industries – communications (such as A.T.&T.) and public utilities; in these industries rate regulation does not permit accumulation of capital through price to the same extent as in the case of non regulated industries.

Against this background of facts, the neoclassical critique must be tested. The facts themselves cannot seriously be questioned. Meticulous scholarship might change the figures by a few billions or an insignificant percentage. The problem is whether change from an aggregate of small-scale individual family-or-ownership-directed enterprise into the conditions indicated by these and like facts entails change or modification of classical economic theory. Since economic theory is in preponderant measure dependent upon assumed motivations, to maintain an unchanged theory must involve assumption that the motivations and possibilities

of action thereon are substantially similar under present conditions as those prevailing before its development.

Peterson feels there is "slender base" for assuming any change. It may properly be suggested that there is even slenderer base for assuming these motivations or possibilities of action thereon have remained the same.

## The shift from "capitalist" control

It would seem today merely whimsy to deny that decision-making control had shifted, away from the "entrepreneur group" of owners who manage, protect and maximize their profits and capital, into the hands of more or less professional corporate administrators. So much so, in fact, that space need not be wasted on extended argument. Any other result (state ownership aside) would be impossible. Save in the diminishing number of enterprises whose founding adventurer or his family still holds an aggregate block of stock sufficient to dislodge a management if they are displeased, stockholders physically cannot, and by law are not permitted, to enter the decision-making process. Further, save in exceptional circumstances (the A.T.&T. may be one such), corporations as a rule do not need and often do not want to have recourse to their stockholders for additional capital. It is maintained, with truth, that the opinions of stockholders do have influence; that stockholders at meetings can raise "pertinent and sometimes embarrassing questions, sometimes with devastating effect"[5] and that they constitute a substantial special public, some of whom at least scrutinize the management. Yet sporadic and only occasionally effective use of this scrutiny does not add up to "control" or anything approaching it. At best, the scrutiny is a variety of post-audit. This is an instance of an old word ("control"), apt in the days of plutocratic 1890s, used by neoclassicists in quite different sense as applied to the discontinuous, occasional, quasi-political corporate processes of corporate government today. Practically its entire content now is that stockholders like to see dividends and market values rise, and, disliking the contrary, complain, seek to find the causes, on extremely rare occasions organize changes, when there is trouble.

In considerable personal experience, the writer has not encountered any situations in which a direct decision to apply (or withhold) capital from a given development, or to enter or refrain from entering a new field, has been decided by stockholders. One need not jump to the conclusion that the administrators of corporations are therefore "irresponsible." But again their responsibility differs in content. They are responsible to the impersonal institutional collective known as "the company"; they are secondarily responsible to the direct desire of stockholders at any given moment to enhance their immediately tangible take or to have losses explained. Stockholders act like an unorganized, usually inert, political constituency. They are a "field of responsibility" – far, indeed, from an entrepreneurial controlling force.

Nor has the situation been materially changed by the practice of endeavoring to make corporation administrators into stock owners through option or other plans. More often than not these plans are endeavors to soften the impact of

income taxes or spread out the high pay of productive years to take care of the administrator's declining years. Rare indeed is the corporate administrator who decides a corporate problem differently because he has ownership of or option to buy a block of his company's shares.

## The immutability of classical economic principles[6]

Less impressive is Peterson's second proposition, namely, that since our economy is dependent on "voluntary self-supporting private enterprise," this fact "precludes departure from its other principal features."[7]

The proposition must be interpreted broadly; as it stands, it is merely bizarre. The American economy was perhaps more dependent on "voluntary self-supporting private enterprise" in the days when monopolies were tolerated than it now is. Private enterprises voluntarily (and enthusiastically) moved into and endeavored to create monopoly situations and to free themselves from competitive restraints (to which we must later pay a little attention). The proposition has to be clarified by adding "under a competitive system." Within limits, addition of the phrase is justified. A powerful structure of antitrust law, Federal trade administration, Department of Justice enforcement, and supporting legal rules in many fields does maintain a version of the competitive system.

Peterson argues, accurately, that under the system "private" (in the sense of non-statist) enterprise must constantly pay attention to obtaining revenue greater than its costs and will seek as great a margin of revenue over costs as can be got. The argument thus runs that the primary object of a corporation must and can only be to maximize its profits, since it is constrained by the forces of competitive conditions. It may not, indeed it cannot, therefore, allow itself luxury of expenditures for social purposes beyond an insignificant margin when profits are healthy. Broadly this is true; but again, the conceptions applying have changed their content.

### *Competition*

Let us begin with "competition."

The first object of competition in the case of large-scale units is to establish that degree of market control, or of equilibrium with other units selling in its markets, so that satisfactory profits may be reaped. One result is the prevalent phenomenon of the "administered price" whose behavior, we are learning, differs considerably from classical patterns. This is not the content of older, classically-described competition. It may, but frequently does not, mean selling in the highest market or buying in the lowest. Sometimes it means pricing to assure entry into, or continued holding of, a particular market sector, though at the time higher prices may be available elsewhere. It may, and very often does, mean low profit or non-profit to increase a market sector, or to fend off some large opponent ambitious to take over. More often its motive is to maintain equilibrium in a market satisfactorily shared with a few colleague-competitors. One may refer to the excellent study by Ralph Cassady, Jr. entitled, *Price Warfare in*

*Business Competition: A Study in Abnormal Competitive Behavior.*[8] The subtitle is accurate. Price competition, beyond the narrowest margin, commonly is abnormal behavior; it breaks out when equilibrium is disturbed or threatened; then it partakes of the nature of an international conflict.

"Competition" at present thus is more often determined by considerations of market power than by those brought to mind by the ancient word. Normally, a state of price equilibrium reasoned satisfactory to all hands is reached, leaving marginal areas only in which the struggle for a customer (or alternatively, the struggle to buy supplies) can be carried on. Most of the time a "live and let live" policy prevails, tacitly, lest there be violation of antitrust laws. The full competitive battery is unmasked only when a newcomer seeks to upset the equilibrium, barging into a reasonably occupied field, or a companion company becomes dangerously aggressive. In great areas, this rarely happens. Factually, if the antitrust laws and state scrutiny were withdrawn, the competitive system would cease to exist in all major lines within a very few years.

Unhappily also for Peterson's argument, a vast sector of the American economy is not, even theoretically, within the classical economic system. Most of wages, all transport, all communication, all utilities, most agricultural products, petroleum and great sectors of metals operate under a system of fixed, not competitive prices, and of regulated monopoly, or of legally-maintained competition. The enterprises involved are, *soi-disant*, private and voluntary; they are actually vast collectives. They are expected to be self-supporting; they are not state-owned; but where not licensed monopolies, they are not in full degree competitive and their markets are in large measure guaranteed; the number of economic forces bearing on them is vastly reduced. Behavior of large-scale enterprise, under these conditions, cannot on the empirical evidence available be fairly assimilated to the "market place behavior" posited by the old theory. There are, it is true, elements of similarity. They are under a degree of restraint, partly by market forces, often by state action. But the impact, the degree, and the results of these restraints have changed.

### *Maximization of profit*

Maximization of profit, it is said, is the prime driving force of corporations now as always in the case of business. Agreed.

Classic (and neoclassical) theory assumes that this fact excludes possibility of significant use of the corporate assets and mechanism for social purposes. Both indeed add that such use not only cannot but should not be made. The corporation's significance is thus limited to that of a profit-seeking unit, having the same motivations and acting in the same way as the classical entrepreneur-businessman. Fundamentally a good deal of this is true. Inaccuracy in using the general concept as guide to assumed motivations and behavior of the corporation arises from the changed state of fact. Maximization of profit in the case of giant corporations not only may, but usually does, mean acting quite differently from the small-scale firm; thus the content of the phrase has changed.

Ably-run corporations think of themselves as perpetual, as dependent on maintaining long-range position and as responsible for meeting market demands (which they hope to increase) for an unlimited future. Their policies thus require and include long-range planning, for periods of five to twenty years ahead. At any given moment, they will sacrifice a portion of immediate profit for long-range position. This takes many forms: toeing up capital to assure future source of supply, foregoing immediate profit for better position in any given market; hazarding resources in experimental operations (some of great size) whose profit potential is undemonstrated, campaigning for a changed tax-position – to take only a few. Of course, they hope the policies adopted will eventually "pay off" in revenue dollars, or in added percentage of market, or otherwise, but the time-dimension is changed. On any given occasion this may mean not buying in the cheapest market and not selling in the dearest; not taking immediate opportunity, but seeking the distant rainbow. And so on *ad infinitum*. Though the profit motive is regnant, it is modified in application, timing and direction by all manner of companion considerations.

Not least among these considerations is a lively appreciation by corporate administrators of the capacity of the state to step in when public dissatisfaction (wholly unconnected with their profits) threatens intrusion through political process. Most really large corporations can, immediately, take measures diminishing costs – for example, transferring, or consolidating company-owned towns, and abandonment of same, dropping overboard unpromising lines of activity, breaking substantial competitors, retiring older employees, but are restrained from doing so by considerations of general welfare or public relations. Clearly they expect their ultimate situation to be better than if they pursued the last dollar of profit. One need not, therefore, deny that a form of "profit maximization" is involved. But the results, market-wise, substantially modify the uncomplicated predictions of classical economists. Elaboration here is impossible: the situations are at once too varied and too fundamental. Enough to say that, when a certain size and degree of market control has been attained, crude following of classically assumed patterns would probably involve the corporation in difficulties with the public, with labor, with the antitrust laws, with legislative and executive authority – though they could make immediate gains. Refusing them is, perhaps, profit maximization – but reinterpreted in the light of modern reality.

Corporate size and concentration is here a powerful, probably a determinative, element. Size extends business decisions from the purely economic into fields of social movement carried on by political action and reaction. An individual trader need think only of himself. A collective trader whose stockholders number hundreds of thousands and whose customers run in millions must think politically as well. Rudimentary political science as well as market economics must be taken into account. Every modern state has assumed responsibility in whole or in part for general economic conditions, and for tolerability of those conditions for most, if not all, its citizens. For a large corporation, the premises on which the state will act and what action it can and is likely to take can never be ruled out.

This suggests that the "instruction in Elementary Economics" contemplated by Peterson must take in much more territory than that envisaged in his paper.[9] It must do more than "take account of the choices, of all people among all goods, of the scarcities of all resources, of all alternative ways of using them," and must endeavor to enlarge the corporation's "worm's eye view" of the forces bearing on it. Factually many, perhaps most, corporate administrators do take elaborate account of these forces, and often maintain expensive staffs for that precise purpose. Most of them realize that at any given moment the "choices of all people" may be determined by monopsonic policies of government (as in defense industries), by power-relations with labor, by currency and credit factors determined by the Federal Reserve Board, even by currents of public thought. The corporate operations may include working out price and wage relationships under the guidance of the Secretary of Labor or even the White House; currents of future need in national defense; plans to supply shifting population; relationships with the Department of Interior or the Department of Agriculture to assure supply; maintenance of regional economic stability in conjunction with local authorities – to name only a few. All of which suggest that the elements of economic "control" posited by classicists and neo-classicists need considerable elaboration.

While necessity of this reappraisal is at least partly a consequence of the size of the corporation, it also results from a modern political-economic factor which now is constant and must never be overlooked. In most developed (and a good many underdeveloped) countries of the world – and certainly in the United States – public opinion and political processes no longer tolerate the results flowing from pursuit of the purely economic and competitive processes to their logical end. The community more often than not prefers continuous employment and stability to the minor price-advantage tossed out by competition. Political action will be invoked against unduly low wages, against undersupply of an essential product, against unemployment, perhaps even against oppressive price fluctuations. In blunt, the state, energized by democratic processes, is always a factor, actual or potential. The "entire range of alternatives on the other side of the market in which it sells and buys"[10] are only some and not necessarily the most determinative elements in the supposed "controls" relied on by classical and neoclassical theory.

This brings us to a brief observation on the progress of the "corporate conscience." (To economists, the phrase is oddly romantic: to lawyers, it is ancient and familiar history and therefore by them better understood. Because a corporation is an artificial legal, and not a human, being, it was held in old common law courts to have no "soul" and therefore no "conscience"; it could not validly take an oath; it was not amenable to moral considerations, and so forth.) Corporations are composed of and managed by men. Each of the administrative group does have a conscience and thus consensus does influence corporate action. In substantial measure, as Peterson rightly says, the "corporate conscience" does have a great deal to do with performing the supply function well, with honesty, upright dealing, and observance of applicable laws. But these same managers have also absorbed the idea that corporations (for better or

worse) are also held responsible by an appreciable sector of opinion for some at least of the social conditions proceeding from their operations – also that, if offensive, these conditions may bring into action the powerful machinery of the state. If corporate managers do not themselves know this, their public relations departments tell them so. The "corporate conscience" may be little more than a lively appreciation of possible consequences either of direct violation of ethics or of social results not tolerable to the community, but it is nonetheless real for all that. Where there is a superior management, its "conscience" transcends this, anticipating rather than remedying deficiencies. Deliberate sacrifice of the firm's long-run prosperity is, to be sure, highly unlikely if not unethical. But one result of a corporate conscience may be the devising of means or even the seeking of governmental or other measures – for example, pension trust funds or even (as in the case of oil) stabilization arrangements – making possible attainment of the desired conditions without that sacrifice. Of this sort of corporate activity there is a very great deal. The law indeed goes farther – it approves and encourages a limited amount of direct corporate philanthropy – though this is less important perhaps than other areas in which the "corporate conscience" has come to be an active force.

## Stockholders derive what influence they have from social-political, not from entrepreneurial, factors

The place of the stockholders as residual recipient of profits deserves a final word. Here classical (and neoclassical) theory reaches romantic heights. It insists on having owner-risk-taking entrepreneurs. The seventeen million stockholders are nominated for the role – no other candidates presenting themselves in the corporate spectrum. Ironically, the facts refuse to write the appropriate script. This writer, believing that control function has shifted away from "ownership," sees little necessity for maintaining the fiction of "owner-entrepreneurship" in the corporate picture, or even substantial reason for having the institution. In any case, willy-nilly, we have not got it. To the contrary, we have, essentially a new form of property.

Desire to discover an "owner-entrepreneurship" or "risk-taking" function in stockholders is basically (I think) an emotional desire to find some functional justification for having stockholders at all. A couple of generations ago, they pulled their weight in the economic boat because they saved, and invested their savings, at hazard of risk and with hope of profit in productive enterprise: in other words, supplied risk capital. They also chose, supervised, contributed to, and controlled management. This justified their existence in classical theory. Solid argument could be made for it. As of today, it is probably true that stockholders have saved (or have inherited past savings). But, as we have noted, these savings no longer are a major source of capital. At best, not over 2 or 3 per cent (often less) of new risk-capital actually entering industrial enterprise in each year is supplied from this source. In overwhelmingly large part, personal savings devoted to buying of stock are used, not to furnish capital to enterprise but to buy out the holdings of some prior stockholder. Nym buys General

Motors stock from Bardolph, who bought it from Pistol, who bought it at 10,000 removes from the heirs of Sir John Falstaff – who did, in fact, invest some money in an original issue of common stock of General Motors at its birth. Nym's purchase is still, quaintly, called "investment" – the word having, as usual, changed its content.

The only facts we know are that Nym's money never did get to General Motors and never will; further, that a half-century having elapsed, Nym's purchase no longer has crucial connection with maintaining General Motors' capacity to acquire new capital by selling new issues of stock. An element of such connection is present – especially in the public utility industry – but so tiny as to be almost invisible. Factually, the stock buying and selling processes carried on through the exchanges have sentimental rather than functional connection with General Motors. Nym's "risk-taking" is the risk of the stock market price fluctuations, completely different from the risk Sir John Falstaff may have taken when he paid good money into the treasury of the nascent motor car enterprise. It is almost though not wholly true that the process is completely independent. The relation is about that of the buyer of a sweepstake ticket to the owner of a race horse whose performance determines the lottery prize – little more. No real reason exists to believe that the entire stock exchange process releases significant amounts of capital for true investment in enterprise, though there may, of course, be a small slop-over margin. Commonly, however, when Nym buys Bardolph's General Motors stock, Bardolph does not finance a new enterprise with the proceeds. He turns around and buys Standard Oil of New Jersey – and so on in millions of transactions.[11]

Dr. Paul Harbrecht has been considering a theory that the stock markets have developed a separate, more or less closed, system of property-holding and exchange, and that this system is essentially independent from the actual productive process. *Prima facie*, there is a good deal of evidence to support the theory. Since, as we have seen, the corporation does not need the stockholders' savings, and the stockholder has no management function – merely vague and occasional quasi-political influence – the classical justification for him as source of capital, or as investment risk-taker, let alone as entrepreneur-manager, simply disappears. He toils not, neither does he spin. He merely expects dividends from capital operation, and an unearned increment of value as the corporation compounds the return on withheld profits ploughed back into the business.

Justification of the stockholder's position, if there is one, therefore, must be found outside classical or neoclassical economic theory. I believe there is such justification though the base is politico-social rather than economic. There are solid values in having men and families attend to their own problems and develop their own lives. That requires that they should have a form of wealth – giving them capacity to choose their ways of consumption, and their manner of living, and power to make their own application of such wealth to their own conception of life. Passive property, like stock, does enable men to do this. Yet it is at once apparent that this justification is valid only in direct ratio as stock is widely distributed among the entire population – ideally, among all of it. As

such distribution goes forward, there is measurable addition to the capacity for self-determination of each holder. Further, as the entire organism of the American economy expands, and as capital values increase, an increasing number of Americans – ideally, all of them – become joint heirs of the system's productivity.

That this distribution is gradually occurring is evident. Thirty years ago only a tiny number of Americans held this form of property (or, for that matter, any income-producing property at all). Today, as we see, 17 millions or more hold some of it directly. Tens of millions more hold it indirectly. The distribution is still not good; 1 per cent of the population of the United States still holds a wholly lop-sided preponderance of it. Yet, quite clearly, progress is being made – though more progress has to be made if the vanished economic justification for such property is to be adequately replaced by its only visible alternative – the social justification.

Simultaneously, one notes, the stock markets, save in vestigial trace, have ceased to be allocators of risk-capital and have become allocators of passive property – irrational, but conceivably capable of development into social institutions no less useful than the great life insurance companies and savings banks.

I do not see, therefore, that Peterson's third point stands up, or indeed that his observations are really relevant to the problem in hand. The supine stockholder is protected by an elaborate system of law – chiefly administrative. Indeed he is (and ought to be) very well satisfied with his position. Until, of course, some revolutionary rises to ask him why he should be permitted to have it, especially if a great many others do not. Then it might be remembered that in great parts of the world, including the fascist as well as communist countries and to some extent semi-socialist Britain, his position has been eliminated overnight (as in Russia) or vastly reduced (as in Nazi Germany) or taken over (as in Britain) or sometimes quietly eroded.

Where the stockholder is maintained in his position – as the United States is endeavoring to do – the fact is not proof of the "deep-rootedness" of "traditional capitalism." Rather it results from tenacious holding of an American ideal of individual capacity to choose his own way of life, and of a system giving individuals enough disposable wealth to implement their choice, and from realization that for these ends this form of distributed wealth, however supine and passive, is a useful if not an essential tool.

## Notes

* Adapted from *Quarterly Journal of Economics*, 79 (1), February 1965, 25–40.

† As example, see, among other recent books, Robin Marris, *The Economic Theory of Managerial Capitalism* (New York: Free Press of Glencoe, 1964).

1 S. Peterson (1965), "Corporate Control and Capitalism," *Quarterly Journal of Economics*, 79 (1), February: 9.

2 *Fortune*, LXIX (July 1964), 179 ff.

3 New York Stock Exchange estimates. These are the results of sampling surveys; a certain caution is indicated. The estimates do, I think, give fair indication of the order of magnitude.

4 See First National City Bank (New York), *Monthly Economic Letter,* July 1964, p. 78.
5 Peterson (1965): 22.
6 Classical economists equate economic laws to laws of physical science: men will always act in the same way under the same conditions. Specifically, they will seek to use their labor and their savings or capital to obtain the greatest available profit.
Let us assume this is true. Even on that assumption, at least two powerful variables at once appear.
What is "available" will be determined by the surrounding structure of law and mores. Interest on loaned money, for example, was not generally available under the medieval system.
Mid-twentieth century development has erected a whole structure of mores and laws precluding or forbidding or endeavoring to prevent results of the competitive system in great areas. The community apparently regards these results as so undesirable (or possibly, so costly) that it is prepared to risk higher prices rather than endure them. It is impossible not to conclude that the available choices are restricted and, even with a self-interest motivation, they have changed.
"Profit" depends on desire. Under medieval mores it was likely as not to include progress toward salvation in the next world; this is why savings were perhaps more often applied to building cathedrals and churches than to constructing profit-making installations.
The argument is made more extensively in my book, *The American Economic Republic* (New York: Harcourt, Brace & World, Inc., 1963).
Neoclassical economics, even if it accepts as immutable the classical premise, must take account of two major variations: (a) that huge institutions are different from individuals and that choices available to individuals within large institutions differ from those available to individual owner-entrepreneurs; and (b) that the mores, politics, and systems of laws built thereon demand results which do not logically flow from the competitive system, certainly as carried on by large institutions, and which shift the application of the self-interest theory.
To the classicist, any interference with his "natural laws" is assumed to invite disaster. In America at least these disasters seem not to have occurred and there is no substantial evidence that they will.
7 Peterson (1965): 9.
8 Occasional Paper No. 11, Bureau of Business and Economic Research, Graduate School of Business Administration, Michigan State University (East Lansing, Mich., 1963).
9 Peterson (1965): 13.
10 Peterson (1965): 13.
11 The argument has been fully made and need not be repeated. See Berle (1962): 433.

## References


Berle, Adolf A. (1962), "Modern Functions of the Corporate System," *Columbia Law Review,* 62 (3): 433–9.

—— (1963), *The American Economic Republic*, New York: Harcourt, Brace & World.

Cassady, Ralph Jr (1963), "Price Warfare in Business Competition: A Study in Abnormal Price Behavior," NBER Occasional Paper no.11, Graduate School of Business Administration, Michigan State University.

Marris, Robin, (1964), *The Economic Theory of Managerial Capitalism,* New York: Free Press of Glencoe.

Peterson, Shorey (1965), "Corporate Control and Capitalism," *Quarterly Journal of Economics*, 79 (1): 1–24.

# Part III

# Perspectives for accounting, law and economics

Lessons from the past

# 8 *Economia Aziendale* and *Betriebswirtschaftslehre* as autonomous sciences of the firm

*Arnaldo Canziani*

## Introduction: the firm as a non-theoretical object in economics before Marshall. The "anti-realistic firm" before and after Alfred Marshall

Looking back over the history of economics, we must acknowledge that, in most cases, the behavior of entrepreneurs or stylized enterprises – rather than the economic nature of the firm as such – attracted the inner speculative attention of scholars.

We range in fact – before and during the Industrial Revolutions – from Necker and others studying public finances, to Bentham and his *felicific* calculus of "social happiness," to the division of labor that starts the growth process according to Adam Smith. Then, from the latter to marginalism, we can read page after page about rightly egoistic butchers, net receipt maximizing monopolists in the field of salubrious mineral mines (Cournot), profit-maximizing railroad monopolists (Dupuit), utility-maximizing neoclassical agents of various kinds (from Jevons to Walras to Pareto to others). In parallel, in German-speaking countries – Menger apart – we can find the fruitful reconstruction of whole economic institutions and agents of the past according to the purely historic approach, as well as to the modern interpretation of markets, capital and industrial groups where firms seem to be mainly stylized characters. Karl Marx could probably have made seminal contributions to the theory of the firm, were he not isolated in the gilded cage of his ideology-driven concepts which brought him to the dead-end of misinterpretations in relation to its economics.

Finally, during this period, almost no author treats the firm as such, nor interprets it in a realistic way, neither in the tradition of (i) A. Smith, D. Ricardo (W. Nassau), and J. S. Mill in Great Britain (thus including the whole galaxy from Petty to Cantillon to Say to others), nor in that of (ii) J. Dupuit, H. H. Gossen, A. A. Cournot, and J. H. von Thünen in continental Europe, to both Menger and Walras (1871–74), to conclude with F. von Wieser and E. Böhm-Bawerk. In those very years, economics was developing and growing and different problems stood out – political ones included. As a result, both scientific subjects, research scope, and value premises were still far from the inner world of firms.

In the last quarter of the nineteenth century, especially in continental Europe, marginalism spread widely in its different approaches and nuances, the so-called "neoclassical tradition." Now, to briefly summarize marginalism in its different forms, we can synthesize its theoretical premises as a theoretical perspective based on: (a) fungible goods, (b) perfect competition, (c) "rational" agents, (d) transparent information, and, in addition, (e) no time-flow, i.e. steady-state. These models were (and are) highly praised as they are indubitable and internally coherent as well as elegant. But they seemed to err on the side of excessive abstractness to both practitioners and applied scholars, who strengthened their critical opinions once Vilfredo Pareto, largely dissatisfied with economics, was the first to abandon in few years this kind of scientific abstraction to turn to sociology.

On the contrary, the English branch led (Jevons apart) to Marshall and the Marshallian firm. Arguments against the English branch could easily stress the never-ending bucolic harmonies of both firms and industries. In fact, the model advocated by the English branch distinguishes between a short and a long period.

In the "short" period the factors are given, and the firm can increase its own output only within the limits of its productive capacity; the firm limits marginal cost and marginal revenue, minimizing the first through the substitution principle; the markets are perfect, the production is homogeneous (in the case of dishomogeneity, the production is sold through special relationships with clients); whenever quasi-rents occur, the entrance of new firms makes production increase and the prices decrease, so that the quasi-rents disappear.

In the "long" period (a mixture of Walras, and Ricardo's rent theory), the well known "representative" firm expands through economies of scale, but its destiny is to decline due to the decrease in average prices: a general renewal of firms is under way, so that the overall distribution of firms according to their age is constant.

John M. Keynes in turn defines the behavior of firms in a way purely targeted to the building of the macro-economic equations representing the aggregate economic variables of his model. For this reason, in chapters 6 and 11 of his *General Theory of Employment, Interest and Money* he presents two or three models – not especially consistent with each other – of the economic transformations of the firm.

In particular, with the same orientation as Marshall, he defines the "marginal efficiency of capital" (*mec*) as the discount rate that equalizes the profitability of capital goods and their present production cost, where (i) the *mec* is never lower than the rate of interest, (ii) it diminishes as long as investments increase, and (iii) investments are increased up to the same rate of interest. Unfortunately, profitability comes from the firm as a whole and is not connected with a single plant, piece of equipment or machinery. This profitability has indeed a very larger variance, and can be (easily) modified by management through administrative decisions in general, alliances and collusions (and cartels) included, as well as increased by economies of scale and market power. In addition, firms make economic evaluations future-to-future, and not only present-to-future.

In this way, it took the end of the Second World War, as well as the meeting with the German stream and the psychological approach – as far as the economics of the firm is concerned – to arrive at a theory of the firm that is jointly

realistic and new: the well known contributions of Machlup, Baumol, Simon, Cohen, Cyert, March (and also Shubik among others), broadly speaking from 1947 to 1968.[1]

We must admit that the substantial revolutions of Machlup, Baumol, Simon and others (J. A. Schumpeter included?) are today largely neglected or even forgotten, as we must also admit that the neoclassical mental construct of the firm as a "black box" has therefore dominated the history of economic thought, and still tends to. This is why "maximizations" are still widespread and surviving concepts – consciously or not – in our mentalities and in so many of our textbooks.

As for this situation which had proved to be unsatisfactory since the end of the nineteenth century, it provoked a reaction especially in continental Europe. This reaction originated in the field of business studies in Germany and Italy in the early 1900s, and reacted against (i) the aforementioned conventional wisdom, (ii) the absence of theories referring themselves to the firm as a whole, i.e. as a dynamic system of actions and resources, and (iii) to the neoclassical tradition (in all its forms) and its lack of realism.

This chapter is devoted to these continental reactions, which in those same years created *Betriebswirtschaftslehre* in German-speaking countries and *Economia Aziendale* in Italy.

## The dissatisfaction in late nineteenth-century continental Europe with conventional wisdom in economics

As we said earlier, the neoclassical, more-than-secular interpretation of the firm – of which we remember today only the maximizing behavior, which interacts with Walrasian markets to magically produce equilibria – is considered as mainly unsatisfactory for scholars in general, and for scholars of business studies in particular. But, as a matter of fact, that framework had been considered as mainly unsatisfactory also by past business studies scholars, since its early appearance, between the end of the nineteenth century and the beginning of the twentieth.

On one side, that interpretation of the firm was only a by-product, as it did not treat its inner economic nature but narrated its behavior according to some (generally false) a priori, consistent with the larger scientific (and sometimes political) aims of the overall theories. Second, some of the previous theories of the firm comprised in the neoclassical tradition were underpinned by an approach that is dominant nowadays, i.e. the mathematical approach, as far from empirical matters as dogmatic in its conclusions (this is the case for Walras, an astronomer; Pareto, a mechanical engineer; Edgeworth, a chemist; Marshall, enrolled in mathematics and a tutor in this subject, and their preference for methods driven by mechanical deductivism).

In continental Europe, the dissatisfaction with the neoclassical theories of the firm – though varying from country to country – emerged immediately, more or less openly. Scholars, if they were trying to study, and to explain as well, the behavior of industrial firms up to the second (or the first) industrial revolution, or to interpret the behavior of banks in those same years – commercial banks,

saving banks, joint-stock ones – found little help (if any) in the usual economic theories of their time. These theories dealt in fact not with firms but with entrepreneurs, intended as perfectly neutral, informed and maximizing single-agents.

These hypotheses (and the above-mentioned others), which were and still are disputable *per se* from the pure theoretical point of view, sounded worst when applied to the realities of that period, and to reality in general: contrasting with historical experience and with facts, they were as false from the descriptive point of view as deceptive from the prescriptive one (as one major goal of scholars in the field of business was precisely to be capable of effectively interpreting phenomena to become capable of prescriptions, according to the distinction by Windelband between *descriptive* and *nomotethical* sciences).

In addition, no further help could be provided by important but largely empirical studies (e.g. Hilferding‘s *Das Finanzkapital*, 1910), as they accumulated large descriptions of phenomena without being able to interpret them from the systematic point of view of the firm. On the contrary, theories must (i) be able to attain an abstract level of generality by making reference to "general" cases (*Idealtypen*) and by generalizing laws and so on, and (ii) not be intended to be immediately practical, nor involved in the nuts and bolts of empiricism. At the same time, theories cannot be so "abstract" as to conform with features that are only partial; in the latter case, they would be based on wrongly or incorrectly selected scientific facts, thus becoming contrary to reality – i.e. shaping fanciful worlds. In addition, they would be so special that they conform only to some industrial or segmental or local specificities.

The methodological controversy of the late eighteenth century in Continental Europe (*Methodenstreit*) had dealt with almost these same problems, and the two streams had been vividly juxtaposed as "theories without facts" vs. "facts without theories." In those very years, the problem for scholars was exactly the same, an everlasting one in truth. On one side, they aim to join real facts with theoretical constructions peculiar to the social sciences, not a construction that is a-critically borrowed from the natural sciences. On the other side, they finally aim to reach consistent syntheses of "scientific facts" related to firms, according to the epistemological revolutions of Mach, Poincaré and others.

Inasmuch as theories satisfy these aims, they inadvertently allow also for orienting practices, at least as general perspectives: as Peter Drucker said, "nothing is more practical than a good theory," while, probably not by chance, the theories of the neoclassical approach clashed (and still clash) with every actual firm, industry, market we could (or can) describe.

Due to these multiple, overlapped and cooperating reasons, those very years – along with pragmatism, and a renewed, critical positivism – resulted in a multivariate, innovative jumble of scientific insights in a number of fields, from the theory of knowledge to applied economics, from innovations in physics to new discoveries in engineering, production and law.

This fervor for innovations was clearly perceived by contemporaries and powerfully stimulated theory and practice, but first in two Continental countries the new insights were synthesized and applied to the business field. Essentially in

Germany and Italy, they turned into unified theories of the firm (and, more generally, of the economic entity).

This peculiar development of business studies was especially due to (i) the development of a late industrial revolution, (ii) the dissatisfaction and stimulation already mentioned, and (iii) the action of seminal innovators such as Nicklisch, Schmalenbach, and Zappa. In fact, other factors have to be mentioned. In particular, (i) the long-standing evolution of accounting studies from Pacioli fifteenth century onwards, (ii) the richness of technical studies in the field of commerce and banking after the sixteenth century, (iii) the revitalization of business studies from the early nineteenth century to culminate with Besta and Penndorf, and (iv) the aforementioned varieties of political economy approaches in those very years, whose contributions concerned rent, value, dynamics and matters other than the firm.

Furthermore, in both countries a new kind of special upper institution was created broadly speaking in the period 1870–1910, the so-called "Upper Schools of Commerce." These schools were devoted to higher training – at graduate and postgraduate level – in the field of business. In Italy, we had the upper schools of Genoa, Naples, Bari and Venice, where studies ranged from accounting to international trade to political economy to law. The Venice school, in particular, recorded the presence and the key role of Francesco Ferrara, a leading economist of his time, and Fabio Besta, who revolutionized accounting during 1880–1917 and transformed it into the "science of control of firms as well as of the wealth they produce." Something similar happened in Germany (and Austria) where, in addition to the influence of Fabio Besta (a scientific leader even in German-speaking countries), and the fast economic development after 1870, the new higher education in the fields of business was developed by *Handelshochschulen*, Higher Schools of Commerce based on the Italian model (Leipzig, 1898; Haachen, Vienna, 1900; Cologne, Frankfurt am Main, 1901; Berlin, 1906; Mannheim, 1907; Munich, 1910, Königsberg, 1915; Nuremberg, 1919).

In those decades (around 1890–1920) and context, business scholars started the hard process of producing realistic interpretations and workable theoretical constructions focused on the firm. This process was fostered also by the 1897–1930 debate on capital among scholars such as Cannan, Fetter, Böhm-Bawerk, Paton, Ricci, Irving Fisher, and by innovative contributions on the special profiles of the firm, the industry or the market (J. B. Clark, 1902; H. J. Davenport, 1913; F. W. Taylor, 1911; A. Marshall, 1911), as well as by the contributions of T. Veblen.

This process – within a few decades, despite (and often against) the reoccurrences or the novelties of political economy – eventually led to the following conclusions. The firm – from Egyptian embalmers and Phoenician traders to present-day Silicon Valley chip-firms – is a special agency in the system, the only one able to produce new wealth, to distribute sectorial and categorical incomes, to reproduce resources. This agency presents the following special features, from (a) to (g). It (a) coordinates production factors; (b) relates to the innovative action of an original agent – the entrepreneur – who has a propensity to risk and looks

at the same time for innovation, self-realization and profit (the goals *a parte subjecti*); and (c) produces "economic goods," i.e. material or immaterial products to be sold on competitive markets (the goods are economic ones because of their function, mainly due to exchanges at the meeting point of supply and demand). These goods are (d) devoted to the satisfaction of human needs, but under the special conditions (very conditions of existence *a parte objecti*) that the revenues reintegrate both period and multi-period costly factors and the investments can be mobilized in time through revenues. In this context, the firm (e) produces wealth in the form of payments, salaries, interests, taxes, and in addition produces new wealth in the form of profits. The latter positive results assure (f) a widespread diffusion and redistribution of wealth among the agents in the system – in the form of categorical incomes, and (g) the so-called civilizing, i.e. the internal and external cultural action it diffuses through design, employees specialization, in-company training, R&D, extra-work incentives, cultural activities *et al.* (civilizing means going back from Perroux to Demaria to Berle to the Italian political scientist of the eighteenth century Gian Domenico Romagnosi).

In accordance with this inner nature of the firm, the actions of the entrepreneurial agent – or, after the "managerial" firm, the actions of the executive core-group – are obviously "subjective": this means that every "maximization," if any, is a subjective one, thus implying all the well known intricacies of information, personal goals and expectations, inter-personal frictions (Machlup, 1946, 1947).

In addition, the firm as a special blend of people, interests, capabilities, right or wrong information and inferences, risk propensities (and limitations and aversions), tends to one further, general blend of goals: sales, profit, growth, self-financing, dividends, salaries, extra-salaries, stock-options, and maybe free-time and public interest as well (Baumol, 1962). As a result, this bulk of conflicting objectives must be consolidated and driven by subjective, "optimally imperfect" decisions.

## The new scientific methodology in continental Europe 1900–1920: the return to a priori[2]

Looking back to the evolution of sciences in the second part of the nineteenth century, they had seen the triumph of two opposite ways of thinking: mathematics and positivism. In fact, these ways are opposite but also similar to each other, both being skeptical with regards to the effective epistemic capacities of human understanding. This is why both take refuge either in the abstract *logos* of mathematics (the exhaustive understanding power of mathematics since Hume and Spinoza) or in the overwhelming importance (and blind attention?) attributed to facts.

Anyway, between the late nineteenth and the early twentieth centuries, it had been clarified that no fact exists without a subjective consideration, especially when scientific facts are concerned. In this way, every science must become "metaphysical" in order to be able to explore the *problems* expressed by facts. Social sciences deserve special treatment and methods, since a pure mathematical approach to them is proved to be partial if not misleading or even deceptive.

This clarification – that might be called an *evolution* in Kuhn's terms – was due to the discontent that every scientific revolution provokes once it wins, dominates, and becomes progressively dogmatic. This discontent reacted, in the case of mathematics, against the excess of blind, aprioristic, and purely abstract deductive reasoning indifferent to facts – or contrary to them; whilst, in the case of positivism, to the excess of empiricism, i.e. the massive a posteriori reasoning and the building of "scientific" theories upon a mess of facts in some cases ordinary, badly defined, fuzzy or even wrong.

Due to this reaction, the massive (and sometimes uncritical) attention given to facts, as well as the treatment of social problems through one-variable equations of the second degree, began to be criticized in the field of social sciences. This opposition was increasingly stronger in the field of business studies. The problem of individuating the inner relationship between facts and theories capable of understanding them was particularly relevant in that field, since it is strictly connected with practices and practical needs, and then must be capable of comprehending these facts in theories of the business firm that are at the same time positive and normative.

In the same years, European positivism was approaching a point of increasing relativism, fostered by the massive a posteriori reasoning in which it was increasingly involved. Its last phase took place in the decades 1900–20, in parallel with the evolution towards phenomenology (Husserl), idealism (Croce, Gentile and Spirito in Italy), subjectivism (Pritchard, Josep, Joachim in England). At the same time, in the USA, the reaction against positivism largely followed, especially in technical studies, Dewey's pragmatism and naturalism.

The evolution of epistemology in those years resulted in an overlapping mixture of three different attitudes, i.e. (a) a blunt anti-positivistic way of thinking, which unified a priori and *theory-building*, (b) a pure positivistic fashion, and (c) an internal critique of positivism.

The opposition to positivism (point (a)) – that could have been carried out in a number of ways – kept the a priori principles or schemes as a premise, as well as the role of the subject in building theories (and facts). The leaders of this movement referred either to Hegel (the neo-Hegelism of Bradley in Britain, Croce in Italy, Royce in the United States) or to Kant (the neo-criticism of Renouvier and others in France, of Cohen, Natorp, Rickert, Windelband in Germany).

Heinrich Nicklisch, the founder of *Betriebswirtschaftslehre* (see next page 114) was inspired by this movement, as his methodology takes from both Kant and Hegel.

Continuing the positivist perspective (point (b)) meant to emphasize, in renewed ways, the role of facts, the impossibility of any a priori framework, and the focus on empirical research (up to a purely empirical degeneration). Since the scholars adopting this perspective were convinced, especially in the field of business studies, that they possessed no other methodology but facts, their approach actually became a sort of *pan-inductivism*, in which the different business disciplines were considered to be nothing else but techniques or *arts*, according to Renaissance semantics. In particular, Eugen Schmalenbach, the

founder of *Privatwirschaftslehre* (second part of fourth section below), belonged to this latter position: in 1912, he wrote *Die Privatwirtschaftslehre als Kunstlehre* (The theory of private economies as an applied art).

During the last evolutionary phase, one internal critique to positivism (point (c)) emerged, aiming at continuing the approach with improvements and further developments. This was especially the case – during 1900–20, or 1905–25 – for two countries of continental Europe, namely Germany and France. They experienced an intense but short-lived "critical positivism." We can remember here, among others, (i) the critical empiricism of Mach, where facts are as important as the relations between them; (ii) the schematism of Boutroux, where scientific laws, being unable to grasp the reality, have no objective value, but only as a reference; and (iii) the conventionalism of Poincaré, where both facts and their relational structure are transformed into science by the scientific research *imposing* its own laws on facts. This internal critique to positivism was relevant for multiple reasons. It allowed the (never ending?) survival of positivism, constituted a junction between positivism and idealism, and proved to be a strong influential factor for European social sciences, law included.

Gino Zappa, the founder of *Economia Aziendale* (see third part of fourth section below), belonged to this tradition where induction and deduction are mixed in the so-called *synthetic* method, which took from many sources, from Croce to J. S. Mill to others (Canziani and Rondo Brovetto, 1992).

## The German schisma: the *Betriebswirtschaftslehre* of Heinrich Nicklisch, the *Privatwirtschaftslere* of Schmalenbach, the business studies 1914–27

### *Nicklisch, or* Betriebswirtschaftslehre *(BWL) as a unified theory of the firm*

Soundly grounded in philosophical studies, throughout his life Heinrich Nicklisch combined his own findings and discoveries in the field of economy with the parallel epistemic revolutions of his time. At ease in deductivistic and organicist systems of thought, due to his culture and attitude he was an institutionalist. As a result, he created the *Betriebswirtschaftslehre*, a new business discipline based on a pure scientific treatment and developed according to a normative approach. The evolution of Nicklisch's thought has been divided as follows:

(a) until 1912: differentiation between *Betriebswirtschaftslehre* and *Nationalökonomie*;
(b) until 1922: definition of the contents of *Betriebswirtschaftslehre* intended as the object of scientific research;
(c) until 1927 and thereafter: the development of the "system" idea and of a systemic approach.

Both *Nationalökonomie* and *Betriebswirtschaftslehre* (point (a)) study the same economic units: families, enterprises, households. The former adopts a *collective*

point of view, the latter the *individual* point of view, i.e. the perspective of single agencies, and the case of single actors inside every unit. Though sharing the same object of knowledge – the economic world – the two sciences do differ in methodology. The goal of *Nationalökonomie* is to investigate the social and economic relationships between agents as a totality. *Betriebswirtschaftslehre* on the contrary first studies single, actual objects, like the firm, whose nature can be both easily recognized in practice and understood in theory, and later their economic life and behavior.

As both the production and consumption of wealth have to be studied through the analysis of the units where these processes typically take place, *Betriebswirtschaftslehre* (point (b)) must study every individual economic unit, be it an enterprise (*Unternehmen*) or a household (*Haushalt*). The enterprise, in particular, is absolutely relevant in the system since it contributes (i) to the satisfaction of human needs, (ii) to the production of new wealth, and (iii) to the circulation of wealth in general. The firm

> is an organic body, a combination of capital and labor which appears externally in defined juridical and economic form. Its actions are largely oriented to its individual interests within a collective economy, anyhow its existence is due to its role in the direct (or indirect) satisfaction of human needs. So, profit as a goal is only a formal profile of the goal-minded efforts of the firm.

Focusing on the firm (point (c)), one relevant problem is the formation of an appropriate accounting system. The firm's contribution to the overall economic system requires measuring its performance over its entire life (*Betriebsleistung*), as well as in single periods of time (financial years so to say). In order to reach this goal and represent capital and income over time, the firm has to be divided into a set of connected and interactive elements (labor, activities, liabilities). Furthermore, it is necessary to define measurement systems that can comprehend both static values (defined in relation with a moment in time, i.e. assets and liabilities) and dynamic values (cost and revenues – income as *die Kraft* – cumulated in intervals called periods).

At the basis of Nicklisch's revolution in the field of business economics – as it happens in almost every revolution – one can find a personal synthesis of many relevant philosophical influences, particularly German philosophers as Kant and Hegel.

As far as Immanuel Kant is concerned, Nicklisch adapted to *Betriebswirtschaftslehre* (and to economic research in general) the ideas of *freedom*, *duty*, *community*. Freedom of action is a natural one for every economic unity acting within the system. The duty of each economic unit is to contribute to the general economic result, as the duty of every human being is to contribute to the life of the community he lives in (this explains also the comprehensive approach of this author, as at the back of his mind he has the concept of *Gemeinschaft*, society in the sense of community, instead of *Gesellschaft*, economic association, or corporation).

It was probably Hegel who passed on to Nicklisch the idea that the spirit (*das Geist*) is given to men, and that activities implying the spirit in some way can be studied only by the spirit itself. In addition, Reason is also action and development, not a merely static Being.

The naturalism of the nineteenth century is also present in Nicklisch's mental constructs: he said, among other things, "everything in nature is commanded by physical laws," and "outside conscience there is the matter. Nature is matter, and it is constituted by force." Later in his scientific life, Nicklisch included naturalism in his own idealistic reinterpretation, talking about "a natural order in the world." No matter if empirical evidence and truth do not match each other: truth is only what is in harmony with the aforementioned "natural order."

Finally, Romanticism too influenced Nicklisch by giving him the concepts of organism and universalism. According to the former, economic units should be studied as single individual entities. According to the latter, the vital connections between parts promote and improve the whole – both within the economic system and for the state.

## ***The relevance of income in Schmalenbach's*** **Privatwirtschaftslere**

Schmalenbach proposed the study of the *Privatwirtschaftslere* (as opposed to Nicklisch's *Betriebswirtschaftslehre*) as an applied science limited to private enterprises and following praxis, i.e. being a *scientia militans*. According to Schmalenbach, "praxis is a sort of client for our science, which has therefore the task of giving it the best possible service." From his perspective, *experience* is the foundation of all scientific knowledge: any element which is not drawn from experience nor is empirically perceptible lies outside knowledge, and outside any positive science as well. Once having denied any influence of philosophy on business economics – as its method cannot be evaluated from a purely theoretical point of view nor judged a priori – Schmalenbach pointed out that "praxis requires an economic science of a non-abstract and non-philosophical a kind, neither looking into the *how* nor the *why* of things." In his opinion, sciences are of two different kinds: theoretical ones, having ideal elements as their object (the theory), and empirical ones, having real elements drawn from the empirical world. *Betriebswirtschaftslehre* and *Privatwirtschaftslere*, fruit of different *Geistes* (intellectual approaches), differ from each other in the same sense: the former is the philosophical science, the latter the practical one, having the goal (and task) of defining optimal behavior and best rules.

Schmalenbach also considered the question of the inner scientific character of his *Privatwirtschaftslere* as irrelevant, since, from his peculiar empirical perspective, the most important thing is the relevance of results. According to Schmalenbach, history says no rule holds general validity, thus we must look for single solutions to the problems of our era: no usefulness of results means no achievement at all.

In Schmalenbach's opinion, the world is dominated by juxtaposed interests, and the basic motive for human action is the *search for individual advantages*. The market is the place of contrasting actions, and prices are the judges of the

quality of the economic performance of human activity. Without capitalism, no motivation for the economic actions of individuals exists, and consequently no room for *Privatwirtschaftslere*. Within markets, economic goods are priced as they incorporate an economic value: this stems both from their scarcity, and from their capability of satisfying needs. This is why the aim of the whole economic activity is to increase the availability of relatively scarce goods.

Moving from these premises, Schmalenbach considered exchange as the starting point of his theory. Exchange is then followed by production, and later by consumption. Though this point is similar to Nicklisch, Schmalenbach actually excluded from the field of analysis every unit driven by consumption (instead of production), such as households (*Haushalte*). Indeed, focusing on enterprises, he concluded his scientific career underlining the importance of proper (objective) measurements of values within the firm through renewed accounting processes (Schmalenbach, 1925).

Seen in this light, accounting takes on the new role of measuring the costs and revenues the enterprise creates and destroys; it is also renewed from the methodological standpoint as knowledge is based on the capacity to observe. The renewed accounting is first of all a method, and an instrument as a consequence: it is a closed system, which allows both the evaluation of the main factors in the economic dynamic of the firm and its self-evaluation. The scope of accounting is therefore the observation, management and control of the dynamics of the single units composing the economic system of the firm. Following this new position, Schmalenbach fostered a revolution in accounting (general accounting, 1908; double-entry and cost-accounting to 1926), according to which the first goal of accounting is not the measurement of capital but of "force" (*die Kraft*). This means placing income as the key parameter for long-term profitability (Hepworth, 1953; Storey, 1960; Boulding, 1962; Schwayder, 1967; Canziani, 1982).

### *The economics of the firm 1910–17: Schär, Dietrich, Leitner, Schmidt, Rieger*

The proposal of a new (branch of) science regarding the firm – that Nicklisch expands to the *plexus* among firms and families – gave rise to a number of debates and disputes, as was already the case in the German *Methodenstreit* (disputes on scientific method). Reinforced by the tradition of autonomy of German universities, as well as by the national preference for clear-cut and theoretically discussed scientific concepts, these debates were concerned with (i) the methodology to be adopted and the role of induction and deduction, (ii) the notion of firm, and (iii) the boundaries of the new scientific field.

As far as the different opinions are regarded, we could place on one side Friederich Schär, according to whom BWL is only a part of economic science, since firms find their *raison d'être* from their relationships with the overall economic system through exchange and commerce. Commerce and negotiations are the bases and the reason for economic action, while they should not be interpreted as profit-seeking activities, but as *the life-source of the economic units*.

Rudolf Dietrich took a quite opposite position, upholding (1914) the autonomy of this field of study, where the firm is the natural focus of the new discipline. The very object of his studies is the internal life of the profit-seeking firm, as well as the relations with other economic units. Fritz Schmidt similarly suggests adopting a two-side approach for the study of the firm (1921): the study of its internal processes and functional relations, and the study of its connections with the general economic system (in his system of thought, induction is subordinated to deduction and is considered as the guideline for any scientific development).

Wilhelm Rieger also devoted his work to the private, profit-seeking firm as the only object of research (1928). According to him, the core problem of the analysis is *value* expressed and measured by money. This is why only the economic units maximizing cash-flows in the long run deserve scientific attention.

Friedrich Leitner sketched a global system (1922), in which the object of analysis is represented by every economic unit, both private and public ones, whether profit-seeking or not. He devoted himself to the study of the private enterprise, treating its general economic dynamics and later dividing it into four main specialist areas: (i) purchasing, (ii) financial management, (iii) communications (from information to transport and distribution), and (iv) sales.

The range of opinions among the German scholars of that time is interesting not only *per se*, but also as they represent usual and recurrent positions in the economics of the firm and offer a range of deep analytical interpretations. Furthermore, a completely new theory of the firm was gradually developed – also thanks to debates and juxtapositions – in those years. In its long and complex process of formation, the following facts were relevant and must be stressed. First, the very center of the analysis was the firm, contrary to every kind of object of the previous one-and-a-half centuries, be they wealth, rent, exploitation, money, taxes, and others. Furthermore, the new discipline emerged from older business analyses, requiring differentiation as well as new methods and tools to investigate the firm as an entity. Finally, the new discipline(s) were confronted with the problems of defining their own object, their methods, their epistemic role, their political space among the pre-existing scientific partitions and developments.

The range of opinions dealt, in particular, with the boundaries of BWL (*Betriebswirtschaftslehre*): is the new science to be concerned only with firms (of every kind and type), or in addition every economic entity inasmuch as they are the very economic agents actually operating within the system?

The same controversy emerged later in Italy, and we can suggest that all those disputes originated under the influence of Husserl. According to him, “phenomena which can be described by laws” first need to define their own boundaries:

> The field of a science is an effectively closed unity; the individuation of the field of truth and the ways of this delimitation cannot depend on our arbitrary judgments. It is a fact that the kingdom of truth is subdivided into fields; investigations have to regard these fields as *objective units* that must be coordinated in sciences. There is a science of numbers, a science of spatial

forms, of human beings, but no autonomous science of prime numbers, of trapezia, of lions or of all these elements taken together.[3]

## The Italian schisma: the *Economia Aziendale* of Gino Zappa from *Tendenze Nuove* (1927) onwards.

*Economia aziendale* as the synthesis between (i) deductivism and inductivism, (ii) accounting, organization and management, and (iii) Nicklisch and Schmalenbach

In Italy, around 1880–1920, two scholars attempted a radical reform of accounting, trying to make it purely scientific in nature. In fact, they served on two opposite fronts and belonged to opposite mental frameworks and constructs:

- Giuseppe Cerboni (1827–1817), who proposed in his *Ragioneria scientifica* (1886–94) to unify accounting and governance, and to apply this new conception to every economic unit, from families to the state;
- Fabio Besta (1845–1922), who proposed in his *La Ragioneria* to treat accounting as the overall science of firms, dealing with accounting, organization and control.

Gino Zappa studied in Milan under professor Bellini, a Cerbonian at large, who later introduced him to Fabio Besta. Zappa became then the best disciple for the most important scholar of the period. The first monograph by Zappa, *Le valutazioni* [Valuations] (1910), was written under the Besta's influence, and was and still is a valuable contribution in the field of evaluation criteria for financial statements including the critique of historical, current, and future values approaches.

Disappointed by the same positivism he had believed in, and stimulated by the intense epistemological debate of those years, after 1911 Zappa started a troubled decade of methodological meditations on theoretical problems related to accounting, such as capital and income, the value of money, the production and distribution of wealth, and was influenced by the ideas of Irving Fisher, the German advances on the matters, and the economic consequences of the First World War, especially the hyperinflation of 1916–20.

As a result of this decade of meditations, Zappa adopted the methodology proposed by "critical positivism," i.e. the joint inductive-deductive or *synthetic* method. According to this approach, scientific facts (i.e. not casual ones, but belonging to classes and series in space and time) are selected by general hypotheses which must be checked, corrected and specified according to the facts themselves. This critical positivism goes back to Bacon and Descartes and brings together empiricism and rationalism up to Kant and beyond. It recalls the last contributions of J. S. Mill (*On the Definition of Economic Policy*) on the mixed a priori method of "induction reasoning," and Ricardo as well. This approach concludes with the active role of the scholar launching and testing hypotheses according to Spencer, Mach, Poincaré, LeRoy and others.

By applying his renewed methodology to the study of both accounting and the firm, Zappa understood (and stated as well), as Schmalenbach did, that *income* is the most important phenomenon in the economy of the firm. Whatever the notion of capital, it has to be related back in some way to income. The firm's income cannot be measured without a deep knowledge, of a scientific nature, of its economy, i.e. of its structure, functioning, dynamics and composition. The renewed method of accounting, then, has to be based on a parallel renewed concept of the firm, or, even better, on a new global science concerning the firm as a whole (*Economia Aziendale*).

Gino Zappa's interpretation of accounting scope and role become nearly dominant in Italy later in time. A strongly innovative conception of firm underlies his accounting revolution, in particular when compared to Besta's theory, which considers the firm as a sum of assets, actions and contracts. The firm was interpreted not as a set of assets and liabilities, nor as a nexus of contracts among the various factors of the production, but as *a set of contracted prices turning themselves* – respectively – *into costs and revenues*. Within the firm's overall life, costs and revenues cross and follow each other, and global income is at the same time the result of these variations and the goal the firm has to achieve.

This renewed concept of the firm, also defined as a "going economic concern," produced a new science: the discipline called *Economia Aziendale*, which means the general theory of the economy of the firm. This theory intended to study with renewed methods the laws of equilibrium and development of every kind of firm (banks and insurance companies *et al.* included). *Economia Aziendale* was proposed and sketched by the official prolusion to the academic year 1926–27 at Venice University: *Tendenze nuove negli studi di Ragioneria* (1927). This contribution, together with the first two parts of his main book, *Il reddito* (*The Income*), published in 1920, are the common fruits of the same revolution.

The notion of income to the firm and the new discipline mutually interact. Income – either global or periodic – has no meaning outside the firm in which it is realized. In addition, it is at the same time its goal and a measure of its efficiency as well as the critical profile of its dynamic economy. Understanding the formation of income requires a sound knowledge of both the economic processes of firms and the effects their dynamics have on present and future costs and revenues. In a word, this requires the new science of *Economia Aziendale*. With its paradigms, this new discipline succeeds in individuating and enlightening new problems by applying its inductive-deductive approach.

Resulting in "a synthesis of accounting, organization science and management," the *Economia Aziendale* has the goal of defining "the dynamic conditions of life" of the *azienda*, which is its field of analysis. Its focus evolved from the firm in an early phase to every economic agency in the later one.

To summarize, the way which brought Zappa from the critical positivism and the revolution of income to the *Economia Aziendale* can be synthesized in six points (see also the sixth section below). First (point 1), sciences – building groups and series of scientific facts – go from the heterogeneous to the homogeneous,

and develop *systems* of facts. According to this systematic, *truth is a quality of the whole system* rather than of its individual parts.

In particular (point 2), applied sciences study both general schemes and specific situations; as the firm is a complex organic unit, the accounting system must represent the whole structure of the connected and interactive economic phenomena, their relationships included. Income (point 3) – the most important fact in the life of the firm – has to be represented in accordance with every other phenomenon of the firm in a systemic way. In this context (point 4), not every administrative fact being an accounting one, and every non-monetary fact having no place in accounting, accounting measurements have to be integrated with the statistical ones. To build a sound integrated system of accounting and statistical information (point 5), the new *Economia Aziendale* is necessary, since method and content cannot be separated. The task of accounting (point 6) is indeed (i) to separate the elements of the organic and unitary life of the firm, (ii) to define values, and (iii) to rebuild the system of the firm according to both wealth and organization (that is why accounting, management and organization enlighten each other and give a synthesis of the dynamics of the firm).

This being true for methodology and accounting, Zappa innovated also from the point of view of *Economia Aziendale* (EA). In the first phase, he agreed with the proposal of Schmalenbach by considering the EA as a branch largely if not exclusively concerned with enterprises, but later in time, he followed the suggestion of Nicklisch by expanding EA to larger boundaries encompassing households, enterprises, and also the state (and its parts). Actually, the methodological premises and the stated principles of accounting-organization-government can be applied to every kind of economic agency.

From this point of view, in conclusion, Zappa's EA resulted in a triple synthesis: (i) of inferential vs. deductive approaches from antiquity to modernity, from Schmalenbach to Nicklisch; (ii) of accounting-organization-governance, turned into a new economic theory, a unitary theory since the economic phenomenon of the firm is unique; and (iii) of Schmalenbach and Nicklisch, taking the income from the first and the global perimeter for EA from the second, and definitively enlarging it.

## The economic nature of the firm

### *Firms and markets according to realism*

According to both *Betriebswirtschaftslehre* and *Economia Aziendale* – in its German or Italian imprint, and within its obviously differentiated ways over decades (Nicklisch, *Die Betriebswirtschaftslehre* [The Economic Theory of the Entity], 1932; Schmalenbach, *Dynamische Bilanz* [Dynamic Accounting], 1925; Zappa, *Il reddito d'impresa* [The Income to the Firm], 1920–29; *Le produzioni nell'economia delle imprese* [Production in the Economics of the Firms], 1955–57; Giannessi, *Corso di Economia Aziendale* [Course of Business Economics], 1965 ff.; P. Onida, *Economia d'azienda* [Business Economics], 2nd edn., 1971; Masini, *Lavoro e risparmio* [Labor

and Saving], 2nd edn., 1980; Azzini, *Economia aziendale* [Business Economics], 1978) – once speaking of "firm" every type of firm has to be included as a matter of fact, be they industrial (including manufacturing), banks, insurance companies and so on. In parallel, the autonomous science dealing with them has the goal – and the duty as well – to work out both descriptive and prescriptive theories for each type. This implies the analytical study of the technical specificity of their economic lives, in any case unified by the following common features: (i) general tendency to economic stability and sustainability; (ii) measurability of their capital- and revenue-values as they are expressed through money; (iii) applicability of such general interpretative categories as income-producing combinations, profitability coordinations (see third part of sixth section below); and (iv) strategic, organizational and managerial processes.

The synthetic approach (i.e. the aforementioned inductive-deductive method) joins the study of sectoral-, industrial-, and firm-specificities to the comprehensiveness of the common interpretative categories. This allows reducing specificities to general rules of growing universality, and enriching and specifying these same rules, giving them concrete forms as well.

As a consequence, empirical studies of this Mach-Poincaré type (neither longitudinal nor merely statistical ones) proved over time the pre-existing scholars' intuitions about firms, markets and competition. In particular, firms differ from each other, since they change in time, grow (or not), confront success, stability or failure. Markets result from the mutable interplay of firms and other agents on the demand side (including families, firms, or the state and its parts). Competition then is both price- and non-price. Each firm confronts competitors and faces the other agents on the demand side (and influences them as well). In this context, each firm may be either a *price-giver* (or *price-proposer*), or a *price-taker* according to the structure and functioning of competition, which includes the existence of dominant firms and cartels.

## *The firm as a system of economic transformation*

According to both *Betriebswirtschaftslehre* and *Economia Aziendale* in their further evolutions, the firm is *the only economic unit (and system as well) which is able to produce incomes and to reproduce capital in a systematic way.* This double capacity also defines its role within the economic system. More specifically, this definition refers (i) to the economic activity orientated *per se* to exchanges within economic dynamics; (ii) with the goal (*a parte subjecti*) and the condition of existence (*a parte objecti*) to produce new incomes, i.e. to achieve a positive difference between revenues and costs (Revenues − Costs > 0); and (iii) with the joint goal (and condition as well) of "mobilizing fixed assets," i.e. to reach the progressive transformation of immobilized multiple-fruition production factors into free cash flows (Broglia, *L'azienda industriale*, 1923).

These three actions – and moments – are subject to different measurements (by accounting) as they are subject to different descriptions within literature, even though carrying out these actions represents a wholeness, a dynamic plexus. Finally, the dynamic firm's economy jointly produces incomes and

reproduces capitals, as cost-prices sum up the contribution of both single-fruition and multiple-fruition factors. In the case where the total costs are covered by income-prices, hence, the economic results express both (a) the net income produced by the economy of the firm, and (b) the cash-flow generated by the transformation of fixed capital.

### *Income-producing combinations, profitability coordinations*

The firm's economy is quantitatively expressed by the accounting synthesis [1]: Revenues − Costs > 0 [1].

In fact, it is far more complicated (see also Biondi, this volume, for further developments). This is why both BWL and EA:

(a) describe it through the economic processes giving life to those quantities;
(b) generalize it from both the economic and quantitative (accounting) point of view.

As far as (a) is concerned, both "costs" and "revenues" are factually the synthesis of multiple sub-systems of costs and revenues. Costs (and revenues) are homogenized by accounting, but they are (highly) dispersed from the contractual point of view as well as in space and time. These sub-systems are summed up and distinguished from each other as far as quantitative determinations are concerned, but they constitute a unity, and a comprehensive system as well, from the economic point of view (point (b)). For instance, costs stand with the (uncertain) goal of obtaining revenues, some costs are substituted for each other; and the whole costs are converted into revenue-hopes. In this way, within the dynamics of markets, through the *continuum* of choices (the aforementioned subjective, "optimally imperfect" choices), income plexuses are molded "in simultaneity and succession," as Zappa said, quoting from J. S. Mill. These (sub-)unities can be defined as income coordinations.

With reference to these income coordinations, the firm is either price-taker or price-giver, according to its various market power on suppliers and clients, to laws and regulations, and to specific managerial choices. In addition, the firm normally negotiates costs before (and in parallel to) the negotiation of revenues, but this process is in some cases inverted (e.g. insurance, re-insurance, contracts, etc.). In short, the firm negotiates the cost-prices of the production factors it chooses, transforms these prices, while also negotiating – afterwards, at the same time, beforehand, "in simultaneity and succession" – revenue-prices. This way it expresses, through its income coordinations, the ways and choices related to its activity in purchasing, producing, and selling. Substantially, the firm is always transforming into costs and revenues the multiple economic processes expressing its different economic activities (purchase-transformation-sales-finance-other).

In this way, the income coordinations constitute the economic parallel of its productive combinations, i.e. the whole set of operations and processes whose interplay produces incomes and renews assets. Within this framework, every

firm is a *productive firm*, since every firm implements productive combinations, creates income coordinations, produces incomes and transforms capitals, banks and insurance and transport included.

Planning, actuating and endlessly changing its "productive combinations," the firm always modifies the related "income coordinations" in space-time, in the endless attempt – provided it is wisely managed – to maximize. But we all know after some sixty years (Machlup, *Marginal Analysis*, 1946, 1947; Baumol, *Expansion*, 1962, *Business Behavior*, 1967; Penrose, *Growth of the Firm*, 1959; Onida, *Economia d'azienda* [Business Economics], 1971) that these maximizations – if any – are intended:

(a) *ex ante*;
(b) in a subjective way (by the entrepreneur, and in most cases today by the core-group, also as effect of a voting process highly influenced by personal interest);
(c) in a partial ignorance of data;
(d) with different personal attitudes as far as risks are concerned, both special and general ones;
(e) within different time-spans;
(f) with different effects as far as the organizational factors are intended, and work (see also (b));
(g) with reference to multiple variables, in some cases conflicting ones, and variously joined, merged, fused, and juxtaposed (economic equilibrium from the income, the financial, the monetary point of view; development; growth, and its rate; market shares; duration of the core-group, stability of the shareholders' majority; organizational equilibria; market power; etc.).

To conclude, one more point has to be briefly added, as both BWL and EA innovated regarding one further field, and their frameworks allow us to clarify it. The firm, as a system of capitals and incomes, is a productive firm inasmuch as it produces new wealth. It is based on capitals to produce incomes, but while producing incomes it reproduces at the same time its capitals, i.e. the assets through which it works. In fact, capitals too are only an instrument: contrary to the *vulgata*, it is not capital (neither tangible nor intangible) which produces income, but the managerial activity. Without management, capital is inactive and useless. Furthermore, bad capabilities managing the same capital produce losses instead of wealth. The following part of this section treats the relevance of management, that of income streams in time, and their conjunction in the firm.

## ***The relevance of management and income***

### *Relevance of management*

Without strategic and operative choices which orientate, set up, and modify the capital, the latter is totally stagnant and unproductive.

The origins and causes of the income-production are – through capital – the whole set of managerial choices. The quality of these choices is the factor able to generate a positive or negative income stream. According to this quality, the profitability of firms may spread on a skewed statistical distribution, i.e. according to a (highly) asymmetric distribution of the single profitability, as told by the well known story of the leading, average or marginal firm. In both expanding and declining periods, the statistical distribution may be *skewed*, even with more asymmetric differences, since, in periods of success or crisis, the ranking of firms by profitability is modified by reason of the different efficiency and effectiveness of managerial choices.

### *Relevance of income-streams in time*

Finally, the effectiveness of capital lies in management, the very origin of income-production. In addition, the value of capital is a function of the future stream of income generated by the management of capital in space-time.

Highly capitalized firms can – due to negative economic situation or tendency, slumps, market troubles, no-holds-barred competition, bad management or even *mala gestio* (fraudulent conduct) – have to confront large, repeated, increasing losses leading to crisis, insolvency or failure. These losses progressively decrease the capital to the point of the impossibility of meeting its own obligations. This leads to insolvency, i.e. to the (compulsory) winding up that nullifies the value of the firm.

On the contrary, poorly capitalized firms can – due to positive trends, market success, low competition and administrative extra-capabilities – gain large, increasing, more-than-average profits. These profits progressively increase the net wealth, allowing for its remuneration as well as the financing of the processes of internal and external growth. And in some cases these profits can lead the firm to acquire a relevant relative position in its industry and also in the overall economic system, increasing its value at the same time.

As a conclusion, the income-streams – different in quantity, quality, time and monetary forms – give value to the capital; in addition, incomes as a result represent the very origin of capital.

### *Something more than discovering prices*

It took decades for many scholars in various nations to build the theoretical construction roughly summarized here, concerned with the economics and management of industrial, banking, insurance and "service" firms over time. But at the end of the story the building was – if still in progress – far richer, more fertile (and more complex) than the aforementioned "black box" and the "two-stage Marshallian firm."

This is surely why Italian and continental scholars paid no attention to *The Nature of the Firm* by Ronald Coase. Before the Second World War, the proposal of this young scholar – today so widespread and accepted after being reintroduced by

O. E. Williamson after 1975 – had attracted almost no attention. Coase spoke of the firm as a "system to discover prices" in connection with the costs of using the market. He argued that – centralizing analyses and negotiations – the firm of optimal dimension lies at the crossing of the implicit cost of using the market, and the costs of both internal organization and imperfect decisions. Now, the relevance of information being given, scholars of that time considered the firm as set up to (i) mix productive factors in a way that individuals would not be capable of arranging, such reaching (partially) optimal dimensions; (ii) put together organizational and managerial capabilities, and thus (iii) be able to innovate systematically; and (iv) share risks. For these reasons, the few readers of that article at that time were surely skeptical about magnifying the function of "price discovery." This latter function was considered a minor one and normally already developed either step by step by artisans as well as by entrepreneurs before the firm, or performed by the already existing firm – by firms in general – forecasting market reactions, and evaluating them month after month.

## Economia Aziendale *as an autonomous economic discipline*

When we speak of the innovative power of the genius, we should never forget that he too fuses dozens of ideas, authors, models, facts and contributions over years of inspiration (and, Thomas A. Edison added, of perspiration).

As far as Zappa is regarded, we should mention among his sources: (i) all the Italian (and also European and Anglo-Saxon) accounting tradition from Cerboni to Besta, from Gomberg to Paton; (ii) a whole galaxy of economists, from the classical ones to the German followers of the historic method, to J. S. Mill, Pareto and the young J. M. Keynes; and (iii) last but not least, the whole group of the *Betriebswirtschaftslehre* scholars who proposed the unitary study of the business field – from families to firms to households – in a purely empirical way (Schmalenbach, 1911–12) or according to an absolute a priori method (Nicklish, 1932).

Some sixty years later, we can point out that Zappa's fundamental contributions focused on the economic nature of the firm through his proposal of a synthetic approach providing a unitary perspective on the firm. In this way, Zappa differentiated himself from both economists and accountants. The former – who studied in the nineteenth century some elements such as capital or market structure or the entrepreneur or the rate of interest – would have studied the "theory of the firm" finally, but in continued neoclassical ways. The latter – business studies apart – would have studied mainly the firm's functions (accounting, finance, marketing), thus atomizing its economic nature. Over the years, the studies by Zappa and his assistants extended also to markets and industries, to look into how markets, exchange structures, and competition could influence the economy of the firm over time. Further fields were explored afterwards: starting from income, firms as entities and the market, the investigations led to the study of consumption and demand, investments, financing, loans and savings, interest rates, bank deposits and stock-exchange behavior.

Approaching the end of his life, Zappa attempted – maybe due to the crisis in economics as well as some heritage from Cerboni, Besta, and the French sociologists of the 1930s – to expand the scientific quest of *Economia Aziendale* to the field of the economic system as a whole, carried out through its components (families, enterprises, public administration) and their interrelationships. The processes of production and consumption were interpreted as special ones in firms, but similar in general within all the economic entities, be they families or public authorities.

In this last evolution, proposed in his unfinished *Le produzioni* (1955–57), the *Economia Aziendale* is seen as a global economic science, an integration of – or almost a substitution for – political economy.

## Conclusion: the role of neglected chapters in the history of economic thought

In a well known passage about the writing of economics, Maffeo Pantaleoni stated that the history of theories cannot be concerned with theories that were demonstrated to be false at a later date. This authoritative claim clashes with the special situation of social sciences, where such a clear-cut judgment is far from being applicable, nor is applicable the experimental proof of some natural sciences such as chemistry or classical physics.

Within economics in particular, due to well known historical, cultural and political reasons, and ideological ones as well, both some "wrong" theories have been developed, studied, diffused and reinforced for decades, whilst others – rightly or not – have left a large stamp on the evolution of political economy and in some cases economic policies too. In parallel, seminal or fruitful contributions were neglected for several reasons, such as the existence of dominant paradigms, usually monolithic ones; an excess of innovation in relation to standard knowledge; the potential upsetting and subversion of accepted "principles"; and peripheral languages.

This being the alternating, asymmetric and largely controversial evolution of economics, the resultant history of economic thought is even more asymmetric, because of the needs of totality, impartiality and order, which naturally clash with both the irregular, un-constant evolution of theories and the capacities (or preferences) of the writers.

This is surely why the history of the economic thought tends to pass directly from J. S. Mill to Marshall to Keynes, thus leaving aside, among others, (i) Wicksell and the Swedish school, (ii) German historicists, Schmoller among them, and (iii) some relevant Austrian and German economists from Böhm-Bawerk to Cassel. As a matter of fact, on the contrary, before the Keynesian revolution, production theory and *Marktformen* (market regimes), cycles and fluctuations, capital and income, and interest theories were the main issues of scientific analyses. After 1936, and especially after the Keynesian *vulgatae* of Hicks in Europe and Hansen in the USA, macroeconomics became in turn the dominating viewpoint of analyses and contributions for some thirty years, and

this evolution overshadowed every different perspective but Marshall and – in some cases – Pigou.

We tended therefore to consider the first part of twentieth century as being a somewhat confused and unimportant period at least in Europe, just a minor segment of the long path from J. S. Mill to Keynes. In addition we considered it as an intermediate phase of political and cultural evolution: the transition between the end of *laissez-faire*, World War I, and its consequences; as well as the transition from positivism (naturalism) to neo-positivism, the Vienna Circle and Wittgenstein.

Nevertheless, should we look at that period with analytical attention, we would discover its deeply interesting features, in that it displays (i) the usual multi-directional confusion of scientific research, (ii) the whole set of premises of every Keynesian, anti-Keynesian or a-Keynesian system of thought, and (iii) last but not least, the initial phase of the economics of the firm in continental Europe (*Betriebswirtschaftslehre* in Germany, *Economia Aziendale* in Italy, as well as *Bedrijfsekonomie* in Holland, *Civiløkonomie* in Denmark, *Företagsekonomi* in Sweden, and so on from Finland to Japan).

In those years (1905–35, broadly speaking), in two European countries – Germany and Italy – accounting and business studies turned into unified theories of the firm, drawing also on the legacy of the whole galaxy of different authors quoted above in the third, fourth and fifth sections. Resulting from a strong methodological turnaround generally based on Kant and Husserl, pragmatism, Mach, Poincaré, Croce and others, these theories – especially *Betriebswirtschaftslehre* in Germany and *Economia Aziendale* in Italy – adopted a research approach infinitely larger than the two-stage firm of the Marshallian tradition.

These two disciplines stated and developed in time, among others, the following principles. First of all, the general scientific method to be adopted is jointly inductive and deductive, and avoids both "theory without facts" as well as "facts without theory." Furthermore, firms in general – banks, insurance companies and others included – were represented as unities of prices in space and time, whose best economic appreciation from the quantitative viewpoint is income. Regardless of industry, juridical form and dimensions, the economics of the firm is a synthetic discipline studying either accounting-production-exchange-finance (*Betriebswirtschaftslehre*), or accounting-management-organization-governance (*Economia Aziendale*). These approaches can be clearly identified by their applied methods such as systemic approach, income analysis, cost-revenue relationships, economic processes, and the formulation of (descriptive and normative) judgments about the sustainability, profitability and development of every kind of firm. Finally, they can be applied to the economy of every agency in the system – from households to firms, from families to the state, i.e. to all economic entities. In this case, they create a science parallel (and perhaps opposed) to political economy. Otherwise, they can be restricted to a clear-cut field of analysis limited to the narrower boundaries of each and every firm.

## Notes

1 Other essays in this book, among them Weinstein and Manfrin, summarize and discuss the more recent partial theories of the firm by Coase, Williamson, and Hart.
2 The third and fourth sections of this chapter constitute a rephrased, enlarged version of some parts of Canziani and Rondo-Brovetto (1992).
3 E. Husserl, 1900, p. 5:

> Das Gebiet einer Wissenschaft ist eine objectiv geschlossene Einheit; es liegt nicht in unser Willkür, wo und wie wir Warheitsgebiete abgrenzen. Objektiv gliedert sich das Reich der Wahrheit in Gebiete; nach dieser objektiven Einheiten müssen sich die Forschungen richten und sich zu Wissenschaften zusammenordnen. Es gibt eine Wissenschaft von den Zahlen, eine Wissenschaft von den Raumgebilden, von den animalischen Wesen usw., nicht aber eigene Wissenschaften von den Primzahlen, den Trapezen, den Löwen oder gar von all dem zusammengenommen.

## Bibliography

Azzini, L. (1978) *Economia d'azienda*, Milano: Giuffrè.
Baumol, W. (1962) "On the Theory of Expansion of the Firm," *American Economic Review*, 52(5): 1078–87.
Baumol, W. (1967) *Business Behavior, Value and Growth*, New York: Harcourt, Brace and World.
Baxter, W. T. and Davidson, S. (1962) (eds.) *Studies in Accounting Theory*, London: Sweet and Maxwell.
Besta, F. (1922) *La Ragioneria*, 3 vols., Milano: Vallardi.
Böhm-Bawerk, E. (1921) *Positive Theorie des Kapitales*, 3rd edn., Jena: Fischer.
Boulding, K. E. (1962) "Economics and Accounting: The Uncongenial Twins," in
Azzini, L. (1978) *Economia d'azienda*, Milano: Giuffré.
Broglia, G. (1923) *L'azienda industriale*, Torino: Mercurio.
Cannan, E. (1897) "What is Capital?" *Economic Journal*, 7(26): 278–84.
Canziani, A. (1982) "Measurements and Calculations in Accounting: A Note on Continental vs. Anglo-Saxon Methodology," *Economia Aziendale*, 1(1): 57–71.
Canziani, A. and Rondo Brovetto, P. (1992) "The Emergence of the Economics of the Firm in Continental Europe during the Twenties: *Betriebswirtschaftslehre* and *Economia Aziendale* as Methodological Revolutions," in S. T. Lowry (ed.) *Perspectives on the History of Economic Thought*, vol. 7, Aldershot: Edward Elgar, 168–91.
Cerboni, G. (1886–94) *La Ragioneria Scientifica e le sue relazioni con le discipline Amministrative e Sociali*, Roma: Loescher.
Clark, J. B. (1899) *The Distribution of Wealth: A Theory of Wages, Interest and Profits*, New York: Macmillan.
Clark, J. B. (1902) *The Control of Trusts*, New York: Macmillan.
Davenport, H. J. (1913) *The Economics of the Enterprise*, New York: Macmillan.
Dietrich, R. (1914) *Betriebs-Wissenschaft*. München: Duncker und Humblot.
Fetter, F. A. (1900) "Recent Discussions on the Capital Concept," *Quarterly Journal of Economics*, 15(1): 1–45.
Fisher, I. (1906) *The Nature of Capital and Income*, New York: Macmillan.
——(1930) *The Theory of Interest*, New York: Macmillan.
Giannessi, E. (1965) *Corso di economia aziendale*, 5 vols., Pisa: Cursi.

Hepworth, S. R. (1953) "Smoothing Periodic Income," *Accounting Review*, 28(1): 32–39.

Hicks, J. R. (1948) *Value and Capital*, 2nd edn., London: Oxford University Press.

Hilferding, R. (1910) *Das Finanzkapital*, Wien: Verlag der Werner Volksbuchhandlung.

Husserl, E. (1900) *Logische Untersuchungen*, Halle: Niemeyer.

Leitner, F. (1922) *Privatwirtschaftslehre der Unternehmung*, 4th edn., Berlin: de Gruyter.

Lowry, S. T. (1992) (ed.) *Perspectives on the History of Economic Thought*, *vol. 7*, Aldershot: Edward Elgar.

Machlup, F. (1946) "Marginal Analysis and Empirical Research," *American Economic Review*, 36(4): 519–54.

——(1947) "Rejoinder to an Antimarginalist," *American Economic Review*, 37(1): 148–54.

Marshall, A. (1919) *Industry and Trade*, London: Macmillan.

Masini, C. (1980) *Lavoro e risparmio*, Turin: UTET.

Menger, C. (1871) *Grundsätze der Volkswirtschaftslehre*, Wien: Wilhelm Braumüller.

Nicklisch, H. (1932) *Die Betriebswirtschaft*, 3 vols., Stuttgart: C.E. Poeschel.

Onida, P. (1971) *Economia d'azienda*, 2nd edn., Torino: UTET.

Paton W. (1922) *Accounting Theory*, New York: Ronald Press.

Penrose, E. T. (1959) *The Theory of the Growth of the Firm*, Oxford: Basil Blackwell.

Ricci, U. (1910) *Il capitale. Saggio di economia teoretica*, Torino: F.lli Bocca.

Rieger W. (1928) *Einführung in die Privatwirtschaftslehre*, Nürnberg: Verlag der Hochschulbuchhandlung.

Schär, J. F. (1921) *Buchhaltung und Bilanz: Auf wirtschaftlicher, rechtlicher und mathematischer Grundlage für Juristen, Ingenieure, Kaufleute und Studierende der Privatwirtschftslehre: mit Anhängen über Bilanzverschleierung und Teuerung, Geldentwertung und Bilanz*, Berlin: Julius Springer.

Schmalenbach, E. (1911–12) "Die Privatwirtschaftslehre als Kunstlehre," *Zeitschrift für Handelwissenschaftliche Forschung*, 6: 304–16.

—(1925) *Dynamische Bilanz*, Leipzig: G. A. Gloeckener.

Schmidt, F. (1921) *Die organische Bilanz im Rahmen der Wirtschaft*, Leipzig: G. A. Gloeckner.

——(1929) *Die organische Tageswerbilanz*, Leipzig: G. A. Gloeckner.

Shwayder, K. (1967) "A Critique of Economic Income as an Accounting Concept," *Abacus*, 8: 23–35.

Storey, R. K. (1960) "Cash Movements and Period Income Determination," *Accounting Review*, 35(3): 449–54.

Taylor, F. W. (1911) *The Principles of Scientific Management*, New York: Harper and Row.

Walras, L. (1874) *Éléments d'économie politique pure*, Lausanne: Corbaz et Cie.

Zappa, G. (1910) *Le valutazioni di bilancio con particolare riguardo alle società per azioni*, Milano: Instituto Editoriale Scientifico.

——(1920–29) *La determinazione del reddito nelle imprese commerciali. I valori di conto in relazione alla formazione dei bilanci*, Torino: Anonima Libraria Italiana.

——(1927) *Tendenze nuove negli studi di ragioneria*, Milano: Instituto Editioriale Scientifico.

——(1955–57) *Le produzione nell'economia delle imprese*, Milano: Giuffrè.

Zappa, G. and Marcantonio, A. (1954) *Ragioneria applicata alle aziende pubbliche*, Milano: Giuffrè.

# 9 The firm between law and economics

## An overview of selected legal-economic scholars of the past

*Thierry Kirat*

### Introduction

The question at the heart of the present work and the contributions discussed herein is that of the enterprise as an entity, thus as an institution distinct from any contractual elements or legal entities (see Berle 1947 [reprinted in this volume] and Manfrin, this volume). All can agree that the capitalist enterprise is something more than a mere series of contracts. Institutionalist economists have the tendency to seek arguments in legal theory that favour the institutional dimension of the enterprise. By doing so, they focus more on legal theory and doctrine than on a deeper analysis of positive law. The thrust of the argument is almost always oriented towards the search for and definition of the substance – contractual or institutional – of the enterprise (see the chapters by Moore and Rebérioux and Weinstein, this volume).

Therefore, the notion of the enterprise is not one-dimensional. In economic analysis, the enterprise is a concept intimately linked to identifiable actions or functions that may consist of the technical combination of production factors (the "black box" canonical conception of the firm), governance of internal transactions (a Transaction Cost Economics concept), the exercise of authority based on ownership of nonhuman assets (Hart and Moore 1990) or the power of employees, owners of human capital, *vis-à-vis* the owners of nonhuman assets (Rajan and Zingales 1998). Such one-dimensional definitions of the concept cannot be found in the legal domain. The basic categories of positive law relative to the production of goods and services of merchants or others are those legal forms concerning mobilization of capital and labour, namely the corporation and the labour contract. The language of law (particularly French law, more generally civil law) certainly recognises the enterprise, but it is an *ad hoc* category, outside the boundaries of core categories, from a perspective of economic and social regulation, in three major areas: work relations, competition and fiscal matters. It is especially on the issue of the integration of salaried employees and the representation of their interests that the question of the enterprise as an "institution" is raised in Europe. As the jurist Georges Ripert acknowledges (1951 [1947]), salaried employees are not stakeholders in the corporation, and it is the corporation itself that exclusively controls the overall development of the corporation,

the freeing-up of supplies, and the relationship between suppliers of capital and managers. More precisely, institutionalisation of the enterprise (since it constitutes more of a process than a state) characterises this movement through which case-law and certain rules pertaining to some specialised law fields (notably fiscal law and social law) have a tendency to give rise to a new legal category, alongside the legal form of a capitalist corporation, geared to guarantee salaried employees a specified legal position in various areas (in particular, the reclassification of salaried employees in case of collective lay-offs and the maintenance of labour contracts in the case of the change of ownership).

Yet, in early-twenty-first-century capitalism, the development of the share value, comparable accounting standards such as IFRS (International Financial Reporting Standards) and legal measures to strengthen the weight of shareholders' interests (including those of employees) in comparison to stakeholders could presage the "return" of the corporation and the reflux of the enterprise. In other words, the enterprise-institution tends to blur when faced with the renewed strength of the central legal form of finance capitalism, that of the capitalist corporation. The latter places the protected interests of the relationship of shareholders and managers back at the centre, to the detriment of employees. Legal theorising and innovation in the category of the "enterprise" have centred on the integration of employees in a new legal category, coupled with the corporation but relatively autonomous. In other words, the enterprise as an institution seems to have lived. However, it continues to pose legal and accounting difficulties (see Biondi, this volume).[1]

A realistic theory of the enterprise should be based on two crucial dimensions: its aspect as an entity or institution, subject to the rules of law and developing its own internal rules; and its role as the operator of the creation and distribution of revenues, and of innovation, thus the economic dynamic. As stated in the first part of this volume by the editors, it is worth characterising the firm-entity as a special economic and financial dynamic and considering it as a *whole* and a dynamic system. Economists, jurists' and accountants' past analyses, not all fitting neatly into the static black box theoretical framework of the enterprise, were examined by Canziani (this volume) in the Italian and German contexts. I would like to extend these reflections through studying the offerings of other legal-economic scholars of the past who provided significant insights into the legal-economic nature of the firm in the period 1920–50. I will demonstrate that the contributions of François Perroux and Georges Ripert in France, as well as those of John Maurice Clark and John Roger Commons in the United States, offer contrasting yet complementary views. I will consider them from the perspective of their contribution to the comprehension of two important aspects previously mentioned: the enterprise as an institution subject to the rules of law, and the enterprise as an actor in the economic dynamic. I will contend that simultaneous apprehension of these two dimensions is a significant analytical and practical issue, and that the institutional-pragmatic vision of Commons is that which best reflects reality. I will reveal that Perroux has advanced an analytical conception of the capitalist enterprise's institutional nature as a motor

of the economic dynamic, and that, from this, Ripert developed a substantial legal vision of it, Clark a "societal" vision and Commons a solid historical, legal and economic vision.

## François Perroux: the enterprise as an institution and the economic dynamic

In French economic thought, François Perroux's work has reflected the desire to place institutions at the heart of the economic dynamic. In reaction to pure theory which considers "the economic process at its most essential . . . drawn from the institutional rubbish," this author estimated that "however simplified, an image of the economic life of men who comprise a the corporation cannot be formulated before registering and characterizing a carefully determined assortment of *institutions*" (Perroux 1951 [1993]: 269).[2]

Three significant aspects characterise Perroux's vision of the enterprise: first, that, in a strict sense, there are only enterprises in capitalist regimes; next, that the enterprise is the key institution of the economic dynamic; and finally, that the economic theory of the enterprise must be intimately tied to that of the creation and distribution of income (and, thus, profit).

### *Capitalism and the enterprise*

Perroux maintained that enterprises only exist in capitalist regimes, the distinguishing feature of capitalism being an "economy of enterprises." The enterprise is the "cardinal institution" of capitalism inasmuch as it is an entity that embodies a technical combination of factors and economic calculations expressed in monetary terms. It assumes that the "calculations and economic wagers by which the enterprise sets up a productive combination [are] attached to the same patrimony," that is, to "the legal and economic collection of values through reference to which the calculations and wagers make sense." (Perroux 1948 [1993]: 18). Capitalism as an economic regime experiences the economic dynamic when entrepreneurs initiate the movement, the rupture of the equilibrium or, in the terms of Schumpeter, of the circular flow. The role of dynamic entrepreneurs is a point stressed by Perroux: contrary to his static counterpart "a slave to past accounting systems" and moved by "the mentality of the *rentier*," the dynamic entrepreneur "places bets for which the economic calculations become the instrument" and "overthrows" static accounting systems (Perroux 1951 [1993]: 20). It is in this context that the dynamic conception of accounting systems discussed by Canziani and Biondi in this volume can be found. As Biondi (2005: 38) explained, the dynamic basis and method are "the accounting way of saying that the firm's economic and monetary process transforms the capital engaged, rather than simply accruing to the total stock of capital. No such thing as permanent capital exists according to the accounting view, but [instead there is simply] cost invested and revenues confronted with actual costs and recovering revenues."

### *The enterprise and the economic dynamic*

In his introduction to the French publication of *Theorie der wirstschaftlichen Entwicklung* [Theory of Economic Development] by Schumpeter, Perroux argued that the Austrian economist's theoretical blueprint should be situated within the framework of his pure abstract economic method. This method, that Perroux proceeds to discuss, stems from the construction of a "representation of a systematically simplified concrete economic reality in order to facilitate the precise comprehension of the characters and the game" (Perroux 1935 [1993]: 79). Thus, it consists of a search for the *essence* of an economic phenomenon rather than its actual and institutional form.

Consequently, in Schumpeter's theory, modifying the circuit and the entrance to the economic dynamic through disruption of the "circular flow" depends on the intervention of the entrepreneur. Yet the vocation of the entrepreneur, an agent of new combinations, is not to name an observable reality.[3] This is an inseparable concept from the pure theory of Schumpeter. Perroux rightly contends that:

> . . . the enterprise and the entrepreneur are unanimously considered as the fundamental province of the mechanism of production, trade and the sharing of a market-based economy. All the authors who have dealt with this recently would accept . . . the formula of W. Sombart according for whom the enterprise and the entrepreneur are the motor of the modern capitalist economy.
>
> (Perroux 1935 [1993]: 138)

Yet a line divides the conceptions of Schumpeter and those of Perroux: for the former, there is no necessary correspondence between the face of the enterprise and the types of economic system (whether capitalism, collectivism, a closed family economy, etc.). For the latter, an enterprise only exists in a capitalist system.

The arguments of Perroux are twofold: first, he specifies that Schumpeter's theory is an abstract theory of the entrepreneur more than of the enterprise, oriented solely towards the question of the conditions for leaving the static mode (circular flow) and entering the dynamic mode. Schumpeter absorbs these conditions into his new functions:[4] (i) the introduction of a new product; (ii) the introduction of a new method of production; (iii) the opening up of a new market opportunity; (iv) the conquest of a new source of supply for raw materials, and the development of a new industrial organization. Therefore:

> The achievement of this new combination is characteristic of the "enterprise" in the sense that Schumpeter employs the word, and the appropriate function of the entrepreneur.
>
> (Perroux 1935 [1993]: 133)

The second section of Perroux's analysis is oriented towards a conception of the entrepreneur that is neither functional nor abstract, but rather constitutes a

historical-sociological and institutional vision of the enterprise. This perception incorporates a psychological component, linking the dynamic with the adventurous spirit of captains of industry, but, above all, it emphasises that profit, the revenue from the capitalist enterprise, is the consequence of networks of power ("*réseaux de force*") and constraints that the state imposes on the activity of the enterprise. The head of a capitalist enterprise benefits from a "delegation of authority" accorded by public bodies, the nature and extent of which we will discuss.

### *The theory of the creation and distribution of revenue*

Thus, Perroux distinguished formal approaches from institutional approaches. From his point of view, a definition of the morphology of an economic process (*à la* Schumpeter) does not excuse an analyst from characterising types of organisations, organisational structures and institutions that Perroux thought existed earlier, as "networks of power" in economic processes.

The shift from a functional approach to an institutional approach was orchestrated by the transition from a purely economic analysis to a historical, sociological and institutional analysis. This is precisely a point on which Perroux constructed his theory of the creation and division of revenues. In effect, he suggested a distinction between three levels of abstraction: first, the abstract level of the approach concerning compensation for the factors of production; next, the intermediate level of "revenues of authority"; and finally the concrete level of "revenues of institutions." (Perroux 1951 [1993]: 267–83).

- The *revenues of factors* are those recognised by Walras's theory of general equilibrium: this separates the distribution from its "institutional matrix"; a similar observation could also be made about the theory of marginal productivity, which considers contributions to the product and the attributions of compensation to be "operated by a mechanism" that "the law of functioning imposes itself on human agents" (Perroux 1966: 964–5). Perroux affirms that the theory of factor revenues, bringing together marginal productivity and equilibrium, is absolutely unrealistic: "in the most liberal, most individualistic, most atomistic of economic societies, the work of production is never accomplished by the simple and spontaneous concourse of pure exchanges. The production unit is never reduced to a series of spontaneously signed contracts." (Perroux 1935 [1993]: 272).
- The *revenues of functions* are those that Perroux finds in Schumpeter's theory of evolution: these are the revenues linked to the functions of authority, innovation, and performance.
- The *revenues of institutions* are those of the capitalist enterprise, capitalist work and capitalist loans.

In that respect, the theory of the revenues of institutions presents certain similarities to Commons's theory of "rationing transactions." The distribution is

studied as a problem of allocating the product rather than the contributions of the product. According to Perroux, this "classic problem of accountability . . . has never been *practically* resolved" (Perroux 1951 [1993]: 275). Perroux claims that the operator of imputations is not – in practice – the market but the head of the enterprise, who evaluates allocations and contributions and divides up the net product according to his own assessments.

> Everything takes place as if, in fact, the head of the enterprise had the delegated authority to appreciate and evaluate the contributions and attributions.
>
> (Perroux 1951 [1993]: 276)

Yet, when state intervention imposes negotiation on groups of enterprise leaders and groups of employees in order to divide up the product, the evaluation of contributions and attributions is clearly determined by a network of power, whereas in a free market economy these networks of power underlie the exchange networks and are hidden by the artifice of the contract between legally equal individuals.

### ***The income of the enterprise: an institutional question***

Perroux does not analyse the contribution of law (of company law or labour law[5]) to the structure of enterprises or the creation and division of the net product. He stresses a more general dimension, the "networks of power," and the "legal-social mould." In that respect, Perroux's argument evolved over time. In 1926, in his doctoral thesis on *The Problem of Profit*, he maintained that net revenues (profits) should be considered in connection with their legal-social mould:

> Thus, we arrive at this conclusion, that concrete revenues cannot be disengaged from their legal-social mould. The concrete revenues are not the product of a factor (Work, Capital, or Land) but must be reattached to an institution. Therefore, there are currently . . . three institutions that comparative law and economics can distinguish, with almost similar characters, from one country to the other: 1. the institution that offers loans and charges interest; 2. the institution offering services (in the form of a free contract); and 3. the institution of the enterprise. These three institutions correspond to three complex revenue streams, each of which have an organic wholeness . . .
>
> (Perroux 1926 [1995]: 337)[6]

Later in his work, he argues that profit is a "revenue of authority" and that the net revenue is different from the surplus of the enterprise according to accounting results (the surplus being defined as the sum of the profits, distributed and not distributed): the accounting result is an artifice of the enterprise's accounting, while the enterprise's revenue is situated in another world, that of the economic dynamics. The functions of creation and authority provide

revenue that is spread over various forms (bonuses to leaders of units of production, directors' compensation, and taxes). The distribution of net revenue is not operated by the market, but by institutions, by networks of power.

After revenues from functions, Perroux finally reaches revenues from institutions (the capitalist enterprise, the capitalist loan, and the capitalist salaried employees). The creation and distribution of these revenues are embedded in economic, social and political institutions and operate within a framework of collective negotiations in which recourse to the private accounting system, as well as to the national accounting system, is inevitable. But, above all, Perroux stresses the fact that the capitalist enterprise, the authority of the head of an enterprise, and profit are intimately linked to the state: the authority of the enterprise's leader is delegated and approved by the state, the other side of which is the imposition of taxes (Perroux 1935 [1993]: 289). The enterprise is an institution that can only be understood within the context of other institutions: it is an institution-entity, endowed with a heritage and an accounting system which allows for monetary calculations and permits the leader of the enterprise to be a "distributor of revenue" while, according to the classic theory of distribution, "it is the market 'anonymous and neutral' which should be [the distributor]" (Perroux 1966: 960).[7]

## *Conclusion*

From Schumpeter, Perroux's theory borrows the thesis of the entrepreneur, motor of the economic dynamic through the use of innovation and recourse to bank credit to break the circular flow. But Perroux goes further than the Austrian economist, whom he reproaches for clinging to functions that, *theoretically*, the enterprise assumes and for not recognising an "observable reality." For Perroux, the enterprise only exists in a capitalist economy: the activity of the enterprise is based on a monetary calculation; it involves the exercise of authority and innovation, and provides revenue, the creation and distribution of which is a significant dimension of capitalism (Perroux 1957 [1991]: 745–61). The enterprise is a key actor in the creation and distribution of revenue, that are institutional processes, and is not the result of an economic equilibrium: they operate within the "networks of power" behind the economic processes and that structure them. These networks of power are created at the macroeconomic level, to divide up revenues. At the microeconomic level, they are formed for the functions of authority approved by the state for heads of enterprises, functions that consist of determining the conditions for calculating contributions to the product and attributions.

Ultimately, while Perroux's theory of the enterprise is intimately linked to his institutional theory of the economic dynamic, it stresses the quasi-political dimension of profit, of firm's income as institutional revenue.[8] Perroux, however, did not consider it useful to situate the question of the enterprise in legal terms, though his commitment to an analysis of the rules governing the creation and distribution of the enterprise's income could have led him there. In his

"dynamic-rules" pairing, Perroux gave priority to the former. It is up to the jurists involved in reflections on the link between the enterprise and capitalism to return to an examination of the second term, the law.

## Georges Ripert's legal theory of the enterprise: beyond contract and corporation

The French jurist Georges Ripert is the author of an important work *Aspects juridiques du capitalism moderne* (Legal Aspects of Modern Capitalism), published in 1947 and reissued shortly after in 1951. While he was unaware of the work of Perroux, his thesis is complementary, not contradictory, to the latter. Likewise, his approach to the question of the enterprise and its relationship with the legal form of the capitalist corporation presents numerous similarities to the theory of enterprise entity defended by Berle (1947, reprinted in this volume).

Ripert devoted himself to analysing the rules and institutions needed by capitalism, particularly for the conduct of economic activity by capitalist enterprises.[9] At the heart of his work is the notion of the enterprise, starting with the observation that it does not constitute a category of positive law, that there is consequently no "law of enterprise." French positive law is familiar with the concepts of contract and the corporation. It is in these ancient legal forms that capitalism creeps in, while preserving the original names and appearance. From this, "[l]egal vocabulary and legal technique obscure the reality" (Ripert 1951 [1947]: 16–17). Under the same legal forms hide such different realities as "the enjoyment of domestic services and the labour contract of an enterprise's workers," and "a corporate contract concluded with two traders united by a brotherly connection and the foundation of a corporation of thousands of shareholders" (Ripert 1951 [1947]: 17).

### *Capitalism and legal concepts*

The history of the capitalist enterprise is a makeshift process involving the rules of the Napoleonic Civil Code, a process of invention of legal rules that capitalism needed in order to expand in the nineteenth century.

According to Ripert, a key question is that of the power to manage capital, thus the capacity to dispose of assets in the course of economic activity, rather than their ownership:

> [I]t is not necessarily the providers of capital who are the masters, but rather those who have the free disposition of capital. The capitalist enterprise is that directed by those who are the control of capital.
>
> (Ripert 1951 [1947]: 17)

Yet civil law recognises only the owner, not the entrepreneur. The history of corporate law is basically a history of a desired opacity in the management of enterprises, the real exercise of which is obscured by a series of "legal fictions."

Ripert sees these legal fictions in corporate law, especially that concerning anonymous corporations. They consist of considering shareholders as the owners of the enterprise and administrators as the agents of the shareholders (Ripert 1951 [1947]: 54), but also treat the corporation as a fictional person, a moral person whose creation depends on the contract between its founders.[10]

Posing the problem of the separation of property and control that Berle and Means analysed in the case of the managerial firm, without resorting to these terms, Ripert sets himself to deconstruct these legal fictions and reveal the non-democratic, discretionary character of control and management of anonymous corporations.[11] The legal forms of the corporation have a formally democratic tinge, to the extent that they decree that sovereignty is exercised by the general assembly of shareholders, with their own office, an agenda for their proceedings, and vote (Ripert 1951 [1947]: 96). As for the administrators, the law considers them agents of the shareholders, instantaneously dismissible by them. Yet, in reality, the shareholders are less sovereign owners than passive creditors of the corporation, and the administrators exercise a power limited only by the *de facto* prerogatives of the managers.[12] Ripert concludes his analysis in these terms:

> Thus the carefully maintained illusion to mask the character of the anonymous corporation is swept away. This capitalist institution would not know how to be democratic.
>
> (Ripert 1951 [1947]: 107)

Reform and regulation of the activity of production therefore depend on the recognition in law of the only effective form that really matters, the enterprise.

### *The enterprise and the law*

Ripert (1951 [1947]: 18) estimates that:

> [t]he law of enterprise has not yet been established. . . . For the moment, it is comprised of borrowings from common contract law and fragmentary legislation on the relationships between the entrepreneur and the holders of capital and the salaried employees.

The issue of the emergence of a law of enterprise is nothing less than the reform of capitalism:

> We do not have a law of enterprise. If one wishes to transform the capitalist regime, it is necessary to create it.
>
> (Ripert 1951 [1947]: 265)

Modern regulation was constrained to invent the concept of the enterprise for its own ends: fiscal law and labour law. In other words, these are the new rules that, unlike common law that ignores it, define a legal notion of the enterprise.

Therefore, since this is a question of special rules, they are not designed to generate a universal category; they define the enterprise for their own purposes. Fiscal law arose so as to consider the enterprise an independent taxpayer, taxes on industrial and commercial profits touching "the income realized by the enterprise" (Ripert 1951 [1947]: 273). To do this, taxation constituted one of the driving forces behind the standardisation of corporate accounting, whereas before "no legal rule determined the manner of drawing up a balance sheet and calculating income" (Ripert 1951 [1947]: 273). Alongside fiscal law, labour law also established the concept of enterprise, despite the fact that it was primarily interested in the establishment itself, the technical unit where work takes place. The enterprise, in the eyes of labour law,[13] constitutes a "society organized hierarchically," according to the terms of the jurists A. Rouast and P. Durand.[14] Thus, for example, the Labour Code specifies that shop discipline constitutes the "law of enterprise" (Book I, Article 22). Corporate committees were established by decree on the 23 February 1945.

Ultimately Ripert estimated, however, that "our current law does not recognize the enterprise" (Ripert 1951 [1947]: 277). Coming back to the definitions of the French economists (Truchy, Baudin, and Nogaro) of the time, he took cognizance of the fact that for the latter the enterprise constitutes a grouping in which a coordination of "human factors and materials of economic activity" takes place.[15] For the jurist Ripert (1951 [1947]: 277), the law – classical law, in any case – "only recognized two ways of co-ordinating its forces: association and work contract" ("*louage [de services]*," literally enjoyment of services). The former is the basis of the corporation's constitution, and the latter is the primary form of labour mobilisation, i.e., the work contract. Consequently, he believed, and this concept has echoes in today's economic debates, that the law cannot view the enterprise as more than a juxtaposition of corporate contracts, loans, and enjoyment of services: "it muddles them to such a degree that there is a certain connection amongst these contracts. It doesn't arrive to merge them" (Ripert 1951 [1947]: 277).

### *Accounting in the enterprise*

It is with regard to the enterprise's capital that Ripert evokes the issue of accounting. From his perspective, the shareholders are neither owners nor managers of corporations through their shares. They are its creditors, meaning that they abandon their right to individual property of their contributions, the other side of which is their right within a corporation to a dividend. As a result, the property of their contributions belongs to the moral person of the corporation. The accounting logic of the balance sheet expresses this ownership: capital, being passive, represents a debt for the enterprise, the other side of which appears in the elements comprising the asset side of the balance sheet. The income appears in the balance of profits (revenues) and losses (expenses). It is not the product of capital, but rather the "product of the enterprise": "It is not born out of the renting of goods accounted in the balance sheet as assets" (Ripert 1951 [1947]: 283).

Ripert challenges the economic schema that makes the capitalist an entrepreneur whose profit is legitimate income deriving from the enterprise's risk-taking. He believes that the model has no empirical consistency and sheds no light on the capitalist enterprise.[16] Starting from the enterprise's accounting system, he argues that the balance of losses and profits, i.e., the net profit, is comprised in the liability side of the balance sheet. This "marks the surplus value of the asset-side" (Ripert 1951 [1947]: 292). Therefore, on the balance sheet, such a net profit consists of "the difference between the value of items in the asset-side and that of the items listed under liabilities. Thus, it is a part of the enterprise's funds. Those who are the owners of the enterprise are, likewise, owners of the profits" (Ripert 1951 [1947]: 292).[17] The entrepreneur's property belongs to "management," which Ripert has shown to be antidemocratic. Consequently, "there is no relationship between the enterprise's profit and the capital invested, so that in a corporation, there is no correlation between the dividend distributed and the nominal value of the share" (Ripert 1951 [1947]: 293).

### *Conclusion*

For Ripert, in conclusion, there is a genuine *de facto* enterprise entity, though legal rules and institutions disguise this with the corporation and the contract. The challenge of Ripert's approach is to define the legal, social and political means of controlling enterprises. For him, liberal French capitalism was able to prosper thanks to the vigour of the legal artifices of the contract and the capitalist corporation, but the time had come to reform capitalism through the legal establishment of a concept of the enterprise. Ripert supplied a profound analysis of the role of law in defining the enterprise, yet offered a rather superficial historical image of the dynamic of capitalism. Instead, he provided a historical perspective of legal forms of economic activity, starting with which he was able to grasp the issues of his times.

## The enterprise as the object of social control: John Maurice Clark[18]

Is the enterprise strictly private? This general question is at the core of J.M. Clark's analysis, and is evident again today in certain approaches to understanding corporate governance. One of the common denominators emphasises the lack of a basic formal notion of the enterprise, in favour of a positive apprehension of legal forms and their impact on the configuration of economic powers, forms of enterprise governance, and mechanisms through which the law organises the representation of interests at play and the forms of social control of enterprises.

### *Social control and non-negotiated rules*

What J.M. Clark meant by "social control" comes back to the role of non-negotiated rules in the course of economic activities: rules custom-designed

specifically for the groupings which are enterprises; rules of legal origin created for and implemented on the occasion of conflicts; legislative rules. By resorting to this notion, in a critical discourse on the individualistic and liberal representation of the market economy, J.M. Clark incorporated in the modalities of social control diverse forms of an adhesion contract that could be considered acts that trigger the application of rules not explicitly part of inter-individual contracts. For this author, the enterprise is an "institution created by law," whose existence and functioning could not have been envisaged outside a legal foundation of rules. The foundation of the corporation is interpreted as a legal innovation by which the legislative power creates a new legal being, possessing prerogatives and powers (Clark 1926: 13).[19] Thus, this author characterises the enterprise as "a game subject to rules," rules taking shape in a variety of institutions: the government at various levels of the federal state (municipalities, districts, states, and the federation) and private institutions (professional organizations, chambers of commerce, unions . . . ). In this context, two points deserve to be underscored: the question of interests, particularly external ones, affected by the modern enterprise; and the role of the contract as an instrument of control.

### *Conflicting interests and legitimacy of private power*

The question of representation of interests and their conflicts is central in Clark's work. Amongst other elements, he wonders about the procedural modalities of resolution of conflicts of interest, and invokes, in that regard, the problem of private economic powers' legitimacy:

> [How should we] resolve the direct contradictions between the conflicting claims between . . . economic groups, not only as to their shares in the national dividend, but also as to matters of power and jurisdiction? For instance, if shop discipline is an exclusive right of the employer, or should the workers have a share in it, and if so, within what limits?
>
> (Clark 1926: 18)

This question comes back to that of the government of enterprises with their dual character. While they are legal beings, they are also human associations pursuing their own objectives, and practising a form of self-government, the content and orientation of which may be subject to social control. The latter is based on questions such as:

> Should labour be admitted to the financial councils and have a voice on questions of production, buying and selling policies, or merely on matters of shop conditions, employment policy, etc.?
>
> (Clark 1926: 70–1)

The legal character *of* the enterprise as well as the legal situation *within* the enterprise can be summarised in the following terms: the business corporation is

a legal being endowed with a moral personality. As such, it is a place where power is exercised. But it is also a reality that individualist thought cannot conceive, a reality shaped by law, an *entity* at the juncture of law and economics. Production activity involves not only the manufacture of goods with a view to satisfying a demand, but also the negotiation of the terms of the division of labour, the definition of personal and real rights that determine to what extent economic activity may turn out to be guilty of "parasitism" and, none the less, remain legal (Clark 1926: 46–7). What is basically at stake is the compatibility between the practices of private enterprises and the public interest. Clark's proposal for reform ultimately comes back to raising the issue of alternative means of social control of enterprises: by the public taking control of the activity as in a collectivist system, by voluntary cooperation between producers and consumers, or by structural changes to private industry. Excluding the first possibility, Clark finds expressions of the second in Europe and envisages reflections of the latter in the United States, in the fields of public utility regulation, in the development of "treaties" between organisations of employees and their employers, and in the "publicisation" of the private contracts. Indeed, he saw signs of its transformation into an instrument of social control in the evolution of the legal doctrine of the contract with respect to the implementation of contractual obligations by the courts:

> The power to enforce carries with it the power to interpret, and the legal conception of a contract includes not merely the agreement itself, but the entire body of law governing its interpretation and enforcement so that social control becomes an integral part of the contract itself.
>
> (Clark 1926: 120)

His analysis of the contract permits him to see therein an instrument of social control, given that, while a contract suggests an agreement, it cannot be reduced to simply that; a contract is an agreement with a view to a mutual exchange enforceable by courts. Consequently, it is a legalised accord to the extent that it creates legal consequences and establishes the possibility of an intervention of the courts with the power to interpret contractual provisions in the case of litigation, also to the extent that the legal regime of contracts is not established in an inter-subjective manner. The agreement's contractual content is also composed of measures of positive law that demand recognition in kind and have a statutory dimension. In other words, the intervention of legislatures in contractual relations is revealed by the power of legal institutions to prescribe part of the content of the contractual agreement, giving credence to the dualistic nature of these relations, both contractual and statutory (Clark 1926: 122).

Finally, it is not by chance that Clark concludes the chapter "Business: Private Right or Public?" of *Social Control of Business* by writing:

> [T]he most powerful individuals in modern industry are corporations. . . . They act, of course, through real persons who are their agents, officials, or

> employees. Presumably a corporation follows its own interests precisely as a single individual would do, but why? If individuals are by nature so devoted to their personal interests that they cannot be trusted to manage their affairs as a collective enterprise . . . , why will not the agents, officials, and employees follow their personal interests at the expense of the corporation? The answer is that they frequently do, and that the maintenance of integrity in our corporate business organizations is one of the fields of public interest and action in modern industry. To this extent, all corporate industry is affected with a public interest.
>
> (Clark 1926: 49–50)

### *Conclusion*

Unlike Ripert, Clark did not question the legal nature of the enterprise as such; he put forward a vision more pragmatic than analytical, based on the social control of enterprises, notably by the law in which he moulded his conceptions in the legal theories critical of his times (that of Wesley Hohfeld and Roscoe Pound). From this perspective, our digression to examine Clark's work is useful in allowing us to reconsider the contractual dimensions of the enterprise. Indeed, displaying two instruments of social control – utility regulation and contract – he considered the latter as a channel through which the enterprise and society enter into contact, and through which the enterprise sees itself as tamed by the law. Therefore, the contract was considered in realistic terms, far removed from what the theory of the firm based on the concept of incomplete contracts proposes today. From this angle, the legal-economic theorisation of Commons is compatible with that of Clark.

## The legal-economic dynamics of the enterprise: John Rodger Commons

Like J.M. Clark, Commons is not concerned with the legal nature of the enterprise as such. Rather than the corporation or the legal form, he is interested in observing the working rules of the institution-enterprise as a going concern, and their transformation through the dynamic of the American capitalist system. While Perroux did not manage to articulate the legal dimension of the enterprise in the dynamic of capitalism, Commons focused on the empirical sense of the enterprise, as an object of social regulation of a legal nature. The articulation of the question of the economic dynamic and of the rules of law applying to enterprises is one significant contribution of Commons.

### *The enterprise as a going concern*

Though Commons has not produced a specific theory of the enterprise, it is, none the less, omnipresent in his work. The enterprise appears as an organisation and as an institution, a place of production and of implementation of rules, all in an

institutional context. The enterprise is a significant case of a going concern, involved in various types of transactions, a basic unit of human interdependencies:

- bargaining transactions, relative to property transfer;
- managerial transactions, relative to conditions of wealth creation;
- rationing transactions, relative to the distribution of the results of the collective action of production.

Commons's transactions are interdependent and their interdependence comprises a whole, the going concern, defined as "a joint expectation of beneficial bargaining, managerial and rationing transactions, kept together by 'working rules'" (Commons 1935: 127). But the enterprise should also conceive itself as an action-oriented trinity. It is, at the same time, a going business, a going plant and a going concern, or, in other words, respectively, a business unit, a productive organisation and an institution.

In defending the institution – understood as collective action in the control, liberation and extension of individual action – as the dominant feature in social life, Commons defines the institutionalist issue with a double thesis (Vanberg 1989):

- individuals always act within the framework of social rules collectively implemented: interactions between individuals are simultaneously conditioned and permitted by the collection of rules, formal and informal, that constitute a society;
- the organisations, understood as units of collective action (corporate actors) are omnipresent in modern society that is, therefore, characterised by a historic process of social organisation, across the development of institutions such as the state, the enterprise, unions, political parties, and so on.

On this basis, Commons offers an original perspective on the enterprise; he conceives it as an organised institution, a sort of coordinating entity underlying the working rules that structure actions and powers. Furthermore, this institution is inscribed in the legal order in the context of shared interests and stakeholders' rights. Commons's category of the enterprise is a particular case of a going concern, a notion forged by the American Supreme Court at the end of the nineteenth century. This associates the idea of an organisation on the move, a dynamic organisation, and that of an institution not simply reducible to the individuals involved, a collective institution.

Commons distinguishes three types of going concerns, characterised by different types of powers, as Dugger (1996: 428) explains:

> The first and most important form of sovereignty is based on the power of the state to exercise violence. Second is the sovereignty of the business concern, which is based on the power of property owners to withhold from others that which they need but not own. This power emerged . . . only after

> the violent power of the state has been limited. Third is the sovereignty of the religious, moral, and cultural concerns, based on the power of opinion.

In a general fashion, the formal and informal working rules deriving from institutions are the sources of rights, obligations and liberties and, at the same time, define the exposure of subjects to the freedoms and rights of other members of the community. In defining what individuals must do, must not do, may do, can do and cannot do, in their interactions with the other, they organise a legitimate domain for authorised transactions and assure the security of expectations without which individuals would not engage in transactions. Consequently, Commons considers the social system to be comprised of the interaction between constitutional and legal rules, the internal rules of economic organizations and general rules of conduct.

### *Public order and private order*

These interactions update a vision of the economic system founded on the articulation of the public order (sovereignty supported by state power) and the private order. Otherwise, Commons's institutionalism maintains that the market economy is inseparable from the legal and political processes that establish it. Thus, it is a social construct. Commons considers the legal-economic nexus to be precisely the sphere where the problem of the social order is constantly posed and resolved. Therefore, the legal rules are interpreted as a special form of working rules, whose enforcement is subject to the control of an authority deriving from the state, and whose interpretation is guaranteed by the courts. These rules have a crucial role in creating the social order through assigning rights, obligations and limits to the exercise of rights and subjection to obligations, through creating powers and limits to their exercise. Consequently, private transactions are authorised transactions that carry rights and duties guaranteed by the sovereign power of the state through legal power, in other words, to authoritative transactions.

### *The enterprise and rules*

If the public order is an important sphere where the transaction rules are defined, every going concern is conceived as a place where the rules are enforced by competent and legitimate authorities, this competence and legitimacy granted by the state in law but also deriving from a specific normative power. Therefore, the rules that structure the functioning of human organisations vary in nature. The rules of common law, those of statute law, shop rules, organisational rules, business codes of conduct, accounting rules, management methods, and rules defined by professional organisations are such different modalities springing from the same movement of the creation of a framework in which transactions take place. Consequently, the enterprise is not an autonomous organisation. It is rather the place where the body of rules stemming from statute law, common law, case law and administration crystallises. The latter

define the framework of transactions between shareholders and managers, employers and employees, the enterprise and clients, etc. But they also appear in the incorporation of the enterprise, i.e. in the form of an "organised collective movement" making the enterprise a singular being (an entity) not only with respect to its members but also with regard to the public order which creates it. At heart, the functioning of the enterprise rests on the exercise of an authority that presupposes the legitimacy of a power of control over the resources and members of the organisation, a legitimacy that allows for the creation of internal rules.

### *The enterprise as an institution*

In analysing the enterprise as a going concern, Commons seeks to grasp both its organisational and its institutional dimensions. The enterprise is an organisation related to the integration of individuals in an entity where they engage in coordinated activity comprised of transactions among the participants, exchange transactions orienting the transfer of property rights, management transactions organising the creation of wealth, and distribution transactions of responsibilities and income amongst participants. As for the institutional dimension, it depicts the enterprise as a collection of relations among participants in a hierarchy, with regulatory powers over the interactions, through a collection of rules defined and enforced by the officers of the enterprise.

The enterprise is an institution characterised by a coordination of "intentions put into effect" based on rules, the purpose of which is the prevention of conflicts between members, as well as an assurance of continued economic activity and the establishment of a framework to resolve conflicts. From this perspective, the origin of the rules basically rests in the decisions of authoritative agencies in matter of dispute resolution. The eminent place occupied by court rulings, particularly that of the Supreme Court in Commons's work, is significant for the role of conflicts of interests in the production of rules.

### *The enterprise and the law*

Commons, unlike Perroux or Ripert, was uninterested in an eventual legal concept of enterprise in itself; for him, it was more pertinent to understand the pragmatic outcomes of regulating enterprises within a given legal-economic system. Uncommitted to a legal concept of the enterprise as such, he devoted himself to the study of rules enforced in organisations, to the modalities of resolving litigation and to the balancing of interests. Indeed, Commons critically evaluated the legal fiction of the enterprise that makes it an artificial, invisible being existing only in the contemplation of the law (Commons 1924: 144 *et seq.*). He focused his attention on the recognition of contradictory interests central to the internal relationships of the enterprise that were brought before the courts. In fact, he was seeking to understand the real manner in which procedures to balance interests at the heart of enterprises and in relationships between enterprises and society operate.

### *Utilities regulation and the accounting problem*

As a member of the team that developed the Wisconsin Public Utility Law of 1907, Commons was involved in the development of railway and public utility regulation in the state of Wisconsin. In that capacity, when defining mechanisms for their regulation that would be reasonable in terms of price, and thus profits, the accounting problem of enterprises was naturally raised. The theory and practice of physical evaluation constituted an important dimension of Commons's work on enterprises affected with a public interest. This concept of physical evaluation "serves as the foundation for setting prices, which are intended by regulators to allow utilities to earn their cost of capital on the depreciated, original cost of their assets, as well as recover their operating costs" (Covaleski 1995). Unlike the neo-classical theory with a purported profit objective, Commons has "advocated the 'rational' deployment of public resources, supported by a complex of such rationalizing techniques as accounting, to solve broad social problems" (Covaleski 1995). The institutional and intellectual context, marked by the recognition of the need to link the accounting question with the balancing of interests (of corporations, shareholders and the public) as the Supreme Court showed in Smyth *versus* Ames (1898),[20] consolidated institutionalists' reflection on the adaptation of an accounting theory of profit based on individualists' economic theories in an emerging corporate economy. Commons's theory of physical valuation expresses the idea that "such notions as accounting profit were flawed by the implicit embedding of organized interests, and that neo-classical economics had taken for granted the very phenomena that needed to be systematically examined: the role of institutional structure, such as economic or accounting information" (Covaleski 1995).

### *Conclusion*

Finally, for Commons, the enterprise is an institution that can only be understood in the context of other institutions, as Perroux argued. From this point of view, the works of the two economists, American and French, are in full agreement. Yet while the French economist saw networks of power in the socio-political contexts of negotiation and distribution of income, Commons saw the imprint of the dynamic of the rules, particularly law, in the definition of the conditions governing the unfolding of transactions within the enterprise.

## General conclusion

The enterprise as a central entity of capitalism is a concept that elicits a dynamic analysis. It is not a stable condition of the price system but rather an actor and subject of the economic dynamic. The accounting theory (see Biondi and Canziani, both in this volume) is useful for understanding the enterprise as an actor in the dynamic – via the emergence and analysis of the enterprise's income – following the model of such economists as Schumpeter and Perroux,

who visualise the enterprise from the perspective of its contribution to the economic dynamic, as an institution of capitalism, a place of monetary calculations. Thus, theorisation on the place of the enterprise in the dynamic of capitalism leaves scarcely any place for law in general, or for the conceptualisation of the economic range of legal forms of the enterprise in particular. On the other hand, the predominantly legal analyses of the enterprise, of its relationships with society, are oriented towards the search for a substantial definition (contract or institutional network) independently of the question of the economic dynamic. One is led to conclude that, finally, there is a trade-off between theorising about the place of the enterprise in the dynamic (without law) and theorising about the legal dimension (which comes back to the static position). On this subject, it seems that Commons's method allows us to escape this impasse though a pragmatic approach to the enterprise as a dynamic institution in a dynamic context.

It is worth emphasising, once again, that the concepts of the enterprise, as well as those of commercial funds or economic activity, are, in practice, notions derived from the dynamic legal-economic framework that also concerns corporations in general – capitalist ones in particular. There is much support for the idea of the enterprise as a collection of human and nonhuman assets, having a permanence that the corporation lacks. The enterprise may persist over time in the form of corporations with limited life spans. However, the perimeter of the enterprise and the interests gathered therein (creditors, employees, managers, clients and suppliers, but also public bodies) is not a constant. It evolves with time, the economic system, legal, social and economic regulation, and accounting standards. The real significant question is that of Commons but also, in very different terms, that of Ripert, the issue of the way in which such a legal-economic framework in general, and positive law and the courts in particular, define the content of the enterprise. Certainly in Europe at least the emergence of the enterprise has corresponded to a period of growing protection of employees' interests in a macro-social welfare state. Does the emergence of a capitalism of shareholders and "financiers" provide a glimpse of the decline of the enterprise-institution and the return of the corporation?

## Notes

1 It appears to be a recurrent debate. Avi-Yonah and Sivan (this volume) show precisely how alternative visions of the alternatives of the enterprise, as an aggregate, as an institutional artifice and as an entity, traverse the history of economic thought, as well as American law.

2 The date within [ ] indicates the year of publication or reissue of the work used. The pagination is that of the work consulted.

3 Stauss (1944; reprinted in this volume) makes the same observation on Schumpeter's theory.

4 "[Schumpeter] refuses purely quantitative variations the right to a place in his dynamic mode and builds it entirely on new functions: the enterprise and credit, and on a new spirit: the 'activism' of the entrepreneur."

5 "[A]s the rule of law in itself interests neither economists nor sociologists nor activists, it is the actual powers that are significant: those of the owner over the means of production, those of labour supply and demand, and those of the governors and the governed." (Perroux 1965 [1993]: 31–2).
6 On this point, see Biondi (2003): 339–47.
7 One may note the remarkable convergence of this analysis with that of Stauss (1944; reprinted in this volume), whom Perroux does not seem to have encountered: the enterprise is an entity that assumes the role of entrepreneur; it is first amongst the participants; the enterprise's accounting is a tool that generates revenue for the enterprise; it is an accounting entity that owns the assets; and it is a bargaining and transacting agency.
8 Economists who will follow Perroux on this point will be determined to demonstrate, like Etienne Antonelli, that the enterprise based on private property and credit is the key institution of the economic dynamic. Antonelli also considered the legal substructure of the capitalist enterprise as a series of fictions from which nothing could be directly deduced about the nature of the dynamic. For example, he distinguished between "legal capital" and "technical capital" of the enterprise. The former involves locked-in assets and the enterprise's working capital. Therefore, for the economic dynamic, only the locked-in productive assets are important, in other words, the "technical capital" (Antonelli 1939, particularly ch. IV).
9 "What is needed in capitalism is a *collection of institutions and rules* that allows for the acquisition and use of capital, ensures that the holding of capital predominates in economic and even political life, and gives the creation and distribution of wealth first place in the spirit of man" (Ripert 1951 [1947]: 15).
10 Contrary to what Rebérioux (2003) asserted, Ripert cannot be straightforwardly associated with a legal institutional theory of the enterprise. He took a critical position with regard to the legal theory of the institution, which makes the enterprise an organic collectivity of labour and capital linked by the idea of a task to be accomplished, according to the terms of the jurist Maurice Hauriou: "The institution is a fashionable but vague expression, since it was created by opposition to the idea of a contract and offers no definite positive characteristics" (Ripert 1951 [1947]: 278). Ripert acknowledges the factual validity of the community character contained in the notion of institution, but he considers that the "difficulty is to find a legal expression within this concept" (Ripert 1951 [1947]: 278).
11 None the less, it appears that Ripert was familiar with the thesis of Berle and Means, since he refers to the work of a French economist who takes them into account, Emile James's *Les formes d'enterprise* (The Forms of the Enterprise). Ripert, however, challenges the idea of shareholders as owners (especially p. 282).
12 "[E]veryone knows that the administrators are designated in advance by those who, according to the common expression, *control the corporation*." (Ripert 1951 [1947]: 284). See Biondi (this volume).
13 From "industrial legislation" as it was called.
14 Rouast and Durand 1951: 89–163.
15 Truchy, *Cours d'économie politique* (Course of Political Economy), 1936, Volume I, p. 1954.
16 Once again, this analysis converges with the critique of Knight's risk theory by Stauss (1944; also in this volume).
17 To be precise, Ripert's "owner" is the one with the power to dispose of and manage the assets, in a way the management or controlling groups of Berle and Means.
18 The following analyses of Clark and Commons are based on previous works in collaboration with Laure Bazzoli (Bazzoli and Kirat 1998).
19 With reference to the disciplinary power of the employer and the subordination of the salaried employees, Clark writes: "[T]his right of corporate management is not by any means an absolute or natural right. It rests on special grants from the state, and is naturally subject to the ultimate power of the state to impose whatever limitations are

necessary to prevent abuses" (Clark 1926: 118). Avi-Yonah (2005) supports the same idea starting from the real entity view.

20 Smyth vs. Ames, 169 U.S. 466 (1898) states that "It cannot . . . be admitted that a railroad corporation maintaining a highway under the authority of the state may fix its rates with a view solely to its own interests, and ignore the rights of the public."

## References


Antonelli, E. (1939), *L'économie pure du capitalisme* [The Pure Economics of Capitalism], Paris: Librairie des sciences politiques et sociales Marcel Rivière & Cie.

Avi-Yonah, R. (2005), "The Cyclical Transformation of the Corporate Form: A Historical Perspective on Corporate Social Responsability," University of Michigan – John M. Olin Center for Law & Economics, Paper No. 05–003.

Bazzoli, L. and Kirat, T. (1998), "L'entreprise et les règles juridiques: une perspective d'économie institutionnelle du droit," *Revue interdisciplinaire d'études juridiques*, 40, 47–77.

Berle, A.A. (1966), "Propriété, production et révolution," in F. Bloch-Lainé and F. Perroux (eds), *L'entreprise et l'économie du XXe siècle* [The Firm and the Economics of Twentieth Century], Paris: PUF, 1966, Tome 3, pp. 681–706. English version: Property, Production and Revolution, *Columbia Law Review*, 65 (1), January 1965: 1–20.

Biondi, Y. (2003) *La nature économique de l'entreprise au croisement des théories économiques et de la comptabilité* [The Economic Nature of the Firm, at the Crossing of Economic Theories and Accounting], unpublished PhD thesis, Universities of Lyon 2 and Brescia.

—— 2005), The Firm as an Entity: Management, Organisation, Accounting, *Università degli Studi di Brescia-Dipartimento di Economia Aziendale*, Paper No. 46, Aug.

Bloch-Lainé, F and Perroux, F. (eds) (1966), *L'entreprise et l'économie du XXe siècle* [The Firm and the Economics of Twentieth Century], Paris: PUF (3 volumes).

Clark, J.M. (1926), *Social Control of Business*, Chicago: The University of Chicago Press (4th impression, 1930).

Commons, J.R. (1924), *Legal Foundations of Capitalism*, New York: Macmillan.

—— (1934), *Institutional Economics. Its Place in Political Economy*: New York, Macmillan.

—— (1935), The problem of correlating law, economics and ethics. Comments on Gény's Méthode d'interprétation et sources en droit privé positif, *Recueil d'études sur les sources du droit en l'honneur de François Gény* [Studies in the Sources of Law. A Tribute to François Gény], Paris: Librairie du recueil Sirey, tome III, pp. 124–44.

Covaleski, M.A. (1995), "The use of accounting information in governmental regulation and public administration: The impact of John R. Commons and early institutional economists," *The Accounting Historians Journal*, June.

Dugger, W.M. (1996) "Sovereignty in Transaction Cost Economics: John R. Commons and Oliver E. Williamson," *Journal of Economic Issues*, 30 (2), 427–32.

Hart, O.D. and Moore, J. (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98 (6), 1119–58.

James, E. (1935) *Les formes d'entreprises* [The Forms of the Firm], *Traité d'économie politique*, Henry Truchy (ed.), tome 3, Paris: Recueil Sirey.

Jeammaud, A., Kirat, T. and Villeval, M.-C. (1996), "Les règles juridiques, l'entreprise et son institutionnalisation: au croisement de l'économie et du droit," *Revue internationale de droit économique*,1, 99–141.

Kirat, T. and Sierra, C. (1999), Economic Space, Institutions and Dynamics: a Revisitation of François Perroux, in M. Bellet and C. Lharmet (eds), *Industry, Space and Competition: the Contribution of Economists of the Past*, London: Edward Elgar.

Perroux, F. (1926), *Le problème du profit* [The Profit Problem] (new edition 1995, Paris: PUG, *Œuvres complètes de François Perroux*).
—— (1935), "La pensée économique de Joseph Schumpeter" [Joseph Schumpeter's Economic Thought], introduction to the French translation of J.A. Schumpeter, *Theorie der wirstschaftlichen Entwicklung*, Paris: Dalloz; collection scientifique d'Economie politique, tome VI (reprinted in F. Perroux, *Œuvres complètes*, vol. IV: *Marx, Schumpeter, Keynes, textes réunis par Renato di Ruzza*, Grenoble: Presses Universitaires de Grenoble, 1993, pp. 61–257).
—— (1948), *Le capitalisme* [The Capitalism], Paris: P.U.F., coll. Que-sais-je.
—— (1951), *Le circuit quasi walrasien à visée historique: les fonctions et les institutions* [The Quasi Walrasian Circuit With Historical Finality: Functions and Institutions], *Economie Appliquée*, tome IV (reprinted in F. Perroux, *Œuvres complètes*, vol. IV: *Marx, Schumpeter, Keynes, textes réunis par Renato di Ruzza*, Grenoble: Presses Universitaires de Grenoble, 1993, pp. 267–83).
—— (1957), *Sur le profit et les progrès économiques* [On Profit and Economic Progress], *Revue de l'action populaire*, 112, novembre (reprinted in F. Perroux, *Œuvres complètes*, Vol. V: *L'économie du XXème siècle*, Grenoble: Presses Universitaires de Grenoble, 1991, pp. 745–61).
—— (1965) *Dialectiques et socialisation* [Dialectics and Socialization], preface to the Edition of K. Marx, *Œuvres*, Economie I, Paris: Gallimard-La Pléïade (reprinted in F. Perroux, *Œuvres complètes*, vol. IV: *Marx, Schumpeter, Keynes, textes réunis par Renato di Ruzza*, Grenoble: Presses Universitaires de Grenoble, 1993, pp. 11–55)
—— (1966), "*Gouvernants de l'entreprise, profit et productivité*" [Firm Management, Profit, and Productivity], in F. Bloch-Lainé and F. Perroux (eds), *L'entreprise et l'économie du XXe siècle* [The Firm and the Economics of Twentieth Century], Paris: PUF, tome 3, 957–91.
Rajan, R. and Zingales, L. (1998), "Power in a Theory of the Firm," *Quarterly Journal of Economics*, May, 87–442.
Rebérioux, A. (2003), *Les figures du collectif en théorie de la firme: confrontation des points de vue juridiques et économiques, communication to the conference, Conventions et institutions: approfondissement théoriques et contributions au débat politique*, Université de Paris X-Nanterre, 11–13 december.
Ripert, G. (1951 [1947]), *Aspects juridiques du capitalisme moderne* [Legal Aspects of Modern Capitalism], Paris: L.G.D.J.
Rouast A. and Durand P. (1951), *Précis de législation industrielle* [Treaty of Industrial Legislation], 4th edition.
Stauss, J.H. (1944), "The Entrepreneur: The Firm," *Journal of Political Economy*, 52 (2), June, 112–27 (reprinted in this volume).
Vanberg, V. (1989), "Carl Menger's Evolutionary and John R. Commons' Collective Action Approach to Institutions: a Comparison," *Review of Political Economy*, 1, 334–60.

# 10 A historical perspective on corporate form and real entity

## Implications for corporate social responsibility*

*Reuven S. Avi-Yonah and Dganit Sivan*

### Introduction

In June 2001, U.N. Secretary General Kofi Annan addressed the U.S. Chamber of Commerce with an impassioned plea for business to "take concerted action against the unparalleled nightmare of AIDS." After discussing the dimensions of the global AIDS crisis, the Secretary General went on to argue that business leaders should get involved in the campaign to stop the spread of AIDS "because AIDS affects business . . . the business community needs to get involved to protect its bottom line . . . there is a happy convergence between what your shareholders pay you for, and what is best for millions of people the world over."[1] Similar calls urging corporations to promote social welfare are repeatedly made. For instance, President Bush and Former Secretary of State Powell have also asked companies to contribute to a global AIDS fund, while former President Clinton also urged companies to attend to social problems (Margolis and Walsh 2003: 268, 269).

The problem with those appeals is that it is hard to show that combating the AIDS crisis in Africa, or any social ill for that matter, will have any discernible impact on the bottom line for shareholders of a local United States corporation. In fact, a recent review of the literature on corporate social responsibility (CSR), the code name for corporations' responsibilities and duties toward multiple constituencies – employees, suppliers, customers, consumers, communities, and society in general –has shown that it is very hard to demonstrate any significant positive correlation between corporate social performance and corporate financial performance (Margolis and Walsh 2003: 278–82).

On the other hand, it is also clear that in many cases corporations, or at least great corporations that are distinguished by their market dominance and the potential political and social power they wield, are in a better position to help human development than either governments or not-for-profit organizations. Corporations are typically smaller and more efficient than unwieldy government bureaucracies and, in the developing world, are also less corrupt. Further, corporations possess greater resources, both financial and technical, than most not-for-profit non-governmental organizations (NGOs).

Moreover, there is evidence that corporations *de facto* tend to engage in corporate social responsibility activities and disclosures. At a minimum, corporations

publicly proclaim to have taken social responsibilities (Williams 1999: 1267–8; Reichert, Webb and Thomas 2000).

Thus, an important question arises: given the facts that corporations do in fact engage in corporate social responsibility activities and are frequently in the best position to help human development, should they be permitted to do so when there is no clear benefit for their shareholders, or should corporations be guided only by a single objective of wealth creation and value maximization? This is a question that has been frequently addressed and debated by academics, starting with the famous Berle–Dodd debate in the 1930s,[2] and continuing in recent corporate social responsibility discussions in the form of the Progressive Corporate Law, the Communitarian Movement, and the Renewed Calls for Corporate Social Accounting and Disclosure.[3] Overwhelmingly, in the last half century, opponents to corporate social responsibility have answered the above question in the negative. From Theodore Levitt's cynical 1958 article on "The Dangers of Social Responsibility" to Milton Friedman's influential New York Times magazine article in 1970, to current writings by Michael Jensen and others, the consensus is that "the business of business is business" and the "social responsibility of business ... [is] to increase its profits" (Chen and Hanson 2004: 37; Branson 2001: 638; Wells 2002: 123). The reasons given are, first, that since management are deploying the shareholders' money, they should concentrate only on one overriding objective, maximizing the shareholders' profits;[4] second, that corporate social responsibility places too much trust in corporate management, and permitting more than one measure of managerial success would enhance the agency cost problem and make it impossible to evaluate managers with any reasonable degree of objectivity (Friedman 1970: 133); and third, by applying the nexus of contracts theory, the dominant legal theory of the corporation, opponents to corporate social responsibility see the corporation as a legal fiction of intertwined sets contracts and as a nonexistent entity that obviously cannot act or form intentions and is incapable of having ethical duties or corporate social responsibilities (Fischel 1982: 1071–4, 1092–4).

And yet, the debate persists not only because most managers, in fact, do engage (or at least appear to engage) in CSR, arguing (in the face of the evidence) that this is in the "long run" benefit of the shareholders (Margolis and Walsh 2003: 270), but also because they are permitted to do so by American Law Institute Principles of Corporate Governance (1994), which state that "[e]ven if corporate profit and shareholder gain are not thereby enhanced, the corporation, in the conduct of its business ... [m]ay devote a reasonable amount of resources to public welfare, humanitarian, educational, and philanthropic purposes" (§ 2.01(b)).[5] This formulation represents a compromise between the wishes of management for maximum freedom and the shareholders' primacy and the nexus of contracts advocates.

This chapter will attempt to shed a new light on this debate by putting it in historical perspective. Historically, the corporation evolved from its origins in Roman law in a series of four major transformations. First, the concept of the corporation as a separate legal person from its owners or members had to be

developed, and this development was only completed with the work of the civil law commentators in the fourteenth century. By the end of the Middle Ages, the membership corporation – a corporation with several members who chose others to succeed them – had legal personality (with the capacity to own property, sue and be sued, and even bear criminal responsibility), unlimited life, and was well established in both civil and common law jurisdictions. The second important step was the shift from non-profit membership corporations to for-profit business corporations, which took place in England and the United States at the end of the eighteenth and beginning of the nineteenth centuries.[6] The third transformation was the shift from closely-held corporations to corporations whose shares are widely held and publicly traded, and with it the rise of limited liability and freedom to incorporate, which took place by the end of the nineteenth century and the beginning of the twentieth.[7] The last major transformation was from corporations doing business in one country to multinational enterprises whose operations span the globe, which began after World War Two and is still going on today.[8]

Each of these four transformations (as well as a smaller, more temporary change which occurred in the United States in the 1980s with the advent of hostile takeovers) was accompanied by changes in the legal conception of the corporation. Remarkably, throughout all these changes spanning two millennia, and every time there was a shift in the role of the corporation, the same three theories of the corporation were brought forward in a cyclical fashion. Those theories include the aggregate theory, which views the corporation as an aggregate of its members or shareholders; the artificial entity theory, which views the corporation as a creature of the state; and the real entity theory, which views the corporation as neither the sum of its owners nor an extension of the state, but as a separate entity controlled by its managers. Moreover, in each shift the real entity theory prevailed, and, for reasons we will discuss below, it was the dominant theory during periods of stability in the relationship between the corporation, the shareholders, and the state.

This chapter is divided into three parts. After this Introduction, we describe in general the historical evolution of the corporation, and in particular we focus on the transformations undergone by the corporate form in the United States during the past two centuries. We show how in each of the transformations undergone by the corporation all three theories tended to arise, but that the real theory ended up as the dominant one. We then go on to draw the normative conclusions and argue that, if indeed the real theory is historically the dominant view of the corporation, it provides a new way of justifying CSR, even when it does not benefit the shareholders and involves problems for which the corporation is not responsible, such as the AIDS crisis.

## The cyclical transformations of the corporate form

The corporation as a legal person separate from its owners is a uniquely Western institution. The corporate form originated in Roman law in its classical

period (the first two centuries AD), was further developed in the Middle Ages in both canon (church) and civil law, and was adopted from civil law by the Anglo-American common law tradition.

In this part, we will describe the transformations undergone by the corporation in the United States during the last two centuries. We will show that, every time there was a shift in the role of the corporation, all three theories (the aggregate, artificial entity, and real entity theories) were brought forward in cyclical fashion. However, every time, the real entity theory prevailed, and it was the dominant theory during periods of stability in the relationship between the corporation, the shareholders, and the state.

The development of the corporation as a legal person, with legal attributes such as owning property and the capacity to sue and be sued, was completed prior to the nineteenth century. The concept of the corporation as a legal person originated in Roman law (Duff 1938).[9] In the period between the classical Roman jurists of the second century and the Commentators of the fourteenth century, this concept gradually evolved. As this evolution proceeded, all three theories of the corporation were brought forward by various legal commentators. Nevertheless, aided by external factors such as the decline of the Holy Roman Empire (which led to the abandonment of the artificial entity theory that the corporation needed imperial permission to exist), the real entity theory emerged as the predominant theory.[10] As we shall see, this pattern of debate among the three theories, followed by the triumph of real entity theory, is typical of subsequent transformations in the role of the corporation, as well.

### *Second transformation: from non-profit to for-profit corporations*

The period between the mid-fourteenth century and the late eighteenth century was one of relative stability in the development of the corporate form. The corporation was established as a membership corporation, i.e., a corporation made up of members who selected their own successors – as do the President and Fellows of Harvard College to this day, for example. As such, a corporation had legal personality with the rights to own property, sue and be sued, and act under a common seal (Clark 1986). Private corporations were used primarily for non-profit purposes (e.g., hospitals and universities), though by the eighteenth century there were also some commercial corporations (e.g., the East India Company) (Blackstone 1765: ch. 18).

During this period, there was reassertion of royal control over corporations; in England and other European countries corporations could only be established by royal charter. Blackstone notes that, although in Roman law corporations could be established without "the prince's consent," "with us in England, the king's consent is absolutely necessary" (Blackstone 1765: 460).[11] None the less, although the king constituted corporations, and the king or other visitors exercised some degree of supervision over them, once established, the corporation (i.e., its members) remained subject to relatively little outside regulation. Therefore, except in extraordinary cases, the real entity view of the corporation

prevailed throughout this period and management (the members) were firmly in control. Corporations, Blackstone notes, were "artificial persons, who may maintain a perpetual succession, and enjoy a kind of legal immortality"; these corporations were "one person in law: as one person, they have one will, which is collected from the sense of the majority of the individuals . . . a person that never dies." This one person then acquires all the rights of corporations, such as perpetual succession, the right to sue and be sued, the right to own property, to have a common seal, to make by-laws, and to be subject to certain criminal liabilities (Blackstone 1765: 455, 456, 463–4).

This situation meant that corporate status was very desirable, especially since the members also enjoyed limited liability for corporate debts.[12] But the English kings were very cautious with granting corporate charters, especially in the case of for-profit enterprises; only corporations that were clearly vested with a public purpose and benefited the public fisc, like the East India and Hudson Bay Companies, received royal approval, and accumulated vast power. This situation, which can be seen as a way of maintaining state control over corporations, meant that the next great shift in the use of corporate form took place in the fledgling United States. There, once the Revolution was over, every state could issue corporate charters. The result was an explosion of charters for commercial enterprises (Angell and Ames 1832).[13]

This proliferation of corporations in the United States was the second great transformation in the role of the corporation in society, from primarily a non-profit to primarily a for-profit enterprise. This profound shift, not surprisingly, led to a revival of the centuries-old debate about the nature of the corporate form and its relationship to the shareholders and to the state. This debate can be seen if we examine the opinions on the subject issued by the first great American jurist, John Marshall. Three of Marshall's opinions, written decades apart, are particularly relevant here: Bank of the United States v. Deveaux (1809), Trustees of Dartmouth College v. Woodward (1819), and Bank of the United States v. Dandridge (1827). These opinions represent the evolution of his thinking on corporations, which moved from the aggregate view (Deveaux) to the artificial entity view (Dartmouth College) to the real entity view (Dandridge).

Deveaux involved an attempt by the state of Georgia to tax the Savannah branch of the Bank of the United States, a corporation established by Congress in 1791, during the early struggles around federalism. The Bank was a membership corporation ("The President, Directors and Company of the Bank of the United States") and all the members were citizens of Pennsylvania. The Bank refused to pay the tax and the State sent its collectors to enforce payment, whereupon the Bank sued the collectors in federal court, claiming diversity jurisdiction. The issue facing the court was whether a corporation made up of members from one state could sue citizens of another state in federal court on diversity grounds. This in turn required deciding between the view that "the individual character of the members is so wholly lost in that of the corporation, that the court cannot take notice of it," and the contrary view that "a corporation

is composed of natural persons," i.e., between the entity (artificial or real) and aggregate views.

Marshall decided in favor of the aggregate view. He stated that the corporation itself, "that mere legal entity," cannot be a citizen or sue in federal court, unless it can be regarded as "a company of individuals" (Deveaux: 86–7). However, since the reasons that led Congress to enact diversity jurisdiction applied to corporations as well, Marshall was inclined to see the controversy as being between the members "suing in their corporate character" and their opponents (Deveaux: 87–8). "The controversy is substantially between aliens, suing by a corporate name, and a citizen . . . in this case the corporate name represents persons who are members of the corporation" (Deveaux: 91). The Court therefore held that federal jurisdiction existed.

Ten years later Marshall was faced with another difficult issue involving corporations. In the famous Dartmouth College case, the state of New Hampshire attempted to alter the charter of Dartmouth College (incorporated as a membership corporation by George III in 1769, under the name of The Trustees of Dartmouth College) by transferring the appointment of trustees to the state, thereby effectively taking it over. The trustees objected, arguing that the charter constituted a contract and that altering it violated the contracts clause of the Constitution.

Marshall held that, because the College was a private corporation, its charter was a contract and was protected by the contracts clause. He addressed the heart of the question – whether the act of incorporation by the state makes it possible for the state to take it over. In frequently quoted language, Marshall held that . . .

> [a] corporation is an artificial being, invisible, intangible, and existing only in contemplation of law. Being the mere creature of law, it possesses only those properties which the charter of its creation confers upon it, either expressly, or as incidental to its very existence.
>
> (Dartmouth College: 636)

This language reflects the artificial entity view of the corporation. But Marshall then went on to note that, having created the corporation, the state may not treat it as a mere extension of itself: "this being does not share in the civil government of the country, unless that be the purpose for which it was created" (Dartmouth College: 636). Even though its object is to promote governmentally approved aims, the corporation is not a mere instrument of the government. Instead, the corporation exists to represent the interest of the founder and his descendants in the aims for which it was founded (Dartmouth College: 642). This interest in the United States is protected by the contracts clause: in this country, "the body corporate, as possessing the whole legal and equitable interest, and completely representing the donors, for the purpose of executing the trust, has rights which are protected by the constitution" (Dartmouth College: 654).

It should be noted that, while Marshall held that the state may not take over a private corporation, even one founded for public ends, the emphasis on the

artificial nature of the corporation left ample room for state regulation via the original charter. Since states were busy granting charters by the hundreds, the Dartmouth opinion enabled the states to regulate corporations, should they wish to do so.

Six years later, Marshall was once more called to opine on the nature of corporations in another case involving the Bank of the United States. The case involved a suit by the Bank on a bond executed by Dandridge, one of its cashiers, in which the defendant argued that the bond had never been approved by the Board of Directors, as required by the charter of incorporation. The key issue was whether the level of evidence required of corporations was higher than that required of individuals, since corporations are incapable of acting not in writing. Justice Story, for the Court, held that no distinction should be made (Bank of the US: 70). Marshall, however, dissented. He argued:

> The corporation being one entire impersonal entity, distinct from the individuals who compose it, must be endowed with a mode of action peculiar to itself, which will always distinguish its transactions from those of its members. This faculty must be exercised according to its own nature. ... This can be done only in writing.
>
> (Bank of the US: 91–2)

The Court's view was the more pragmatic one, but Marshall's view was more consistent with the real entity view of the corporation as distinct from its members, individually or collectively.

How can one explain the shift in Marshall's view of the corporation from aggregate (Deveaux) to artificial (Dartmouth College) to real (Dandridge)? In part, this stems from the circumstances of these particular cases. In Deveaux, Marshall wanted to confer diversity jurisdiction to protect a federal institution (he was after all a Federalist), and the only way to do so was to look through the corporation to its members. In Dartmouth College, the issue involved the relationship between private corporations and the state, and Marshall emphasized the role of the state in creating the corporation, while placing clear limits on its ability to regulate corporations thereafter. These limits were required as the result of the proliferation of corporations, especially for-profit business corporations, since otherwise the state would be able to take over purely private businesses. The result in Dartmouth College favored in practice the real entity view, because once a private corporation was created, it could no longer be taken over or perhaps even be overly regulated by the state. Thus, it may not be surprising that by the time he came to write his Dandridge dissent Marshall took the real entity view, even though it contradicted his opinion in Deveaux (which is not mentioned).

Two important legal developments during the same period strengthened the real entity view and weakened the aggregate and artificial entity views of the corporation: the rise of limited liability and the spread of general incorporation laws. Limited liability weakened the aggregate view, and general incorporation

weakened the artificial entity view.First, limited liability: in the United States, many states adopted limited liability in the 1830s, and by 1840 limited liability was established in most of the states (Blumberg 1993: 11–12).[14] Limited liability, in turn, led to a decline in the emphasis on the aggregate theory, because the aggregate view of corporations tends to reduce the distinction between the corporation and its members or shareholders, which is at the heart of limited liability (Blair 2003: 388–90; Hansmann and Kraakman 2000: 393).

The decline of the aggregate view can clearly be seen in two cases from the period 1839–44, in which the Supreme Court repudiated Marshall's opinion in Deveaux. In Bank of Augusta v. Earle (1839), the Court held that a corporation incorporated by Georgia may execute a valid contract in Alabama on comity grounds, but it rejected the argument that Alabama was required to accept the contract on the basis of the privileges and immunities clause applied directly to the corporation's members (as required by the aggregate view), stating that Deveaux had never been extended that far. Chief Justice Taney emphasized that he rejected the aggregate view because of its implications for limited liability, as well as its implications for state regulation of the corporations operating in it:

> The result of this [aggregate view] would be to make a corporation a mere partnership in business, in which each stockholder would be liable to the whole extent of his property for the debts of the corporation. . . . Besides, it would deprive every state of all control over the extent of corporate franchises proper to be granted in the state.
>
> (Bank of Augusta: 586–7)

In Louisville, Cincinnati, & Charleston Railroad Co. v. Letson, decided in 1844, the Court explicitly limited Deveaux to its facts, holding that diversity jurisdiction may arise even when some of the members of a defendant corporation are citizens of the same state as the plaintiff (Louisville: 559). The Court stated that the Deveaux results "have never been satisfactory to the bar" and that a corporation "seems to us to be a person, although an artificial one, inhabiting and belonging to that state [of incorporation], and therefore entitled, for purposes of suing and being sued, to be deemed a citizen of that state" (Louisville: 555).[15]

This result was required by the proliferation of business corporations having many shareholders in many states, as opposed to the membership corporations of Marshall's early days. The separation of management from ownership, and the rise of limited liability, rendered the aggregate view implausible.[16]

Second, general incorporation: laws were passed in all the states permitting anyone to form a corporation on payment of a fee, without permission by the state legislature.[17] This democratizing move meant that the artificial entity theory, under which the corporation derives its powers from the state, lost most of its appeal because the state was only vestigially involved in creating corporations. Instead, corporations were viewed as separate from both their shareholders and the state, and the real entity view reigned supreme.[18]

### *Third transformation: from closely-held to widely-held corporations*

The situation between the 1820s and the end of the Civil War was the proliferation of for-profit corporations, incorporated under general incorporation laws with minimal interference by the state, and whose shareholders enjoyed limited liability. Those shareholders were, however, relatively limited in number, as few corporations before 1865 required massive amounts of capital, and most were small, closely-held enterprises. This enabled the Civil War tax on corporate income to be imposed directly on the shareholders of corporations (Avi-Yonah 2004: 1212–15).

This state of affairs began to change with the advent of the railroads, followed by the steel and oil companies. With the rise of large corporate enterprises, massive amounts of capital were required, and between 1865 and the 1890s the widely held, publicly traded, non-owner-managed enterprises gradually became the norm for U.S. business activities. This was followed from 1890 to 1906 by a wave of consolidation that left several important business areas dominated by oligopolies run by the "robber barons."[19]

The shift from small, closely-held enterprises to massive, publicly-held ones once again necessitated a re-examination of the corporate form, and again all three theories of the corporation appeared. A classic example of the aggregate view is the Santa Clara case, ultimately decided by the Supreme Court in 1886. This case is famous for Chief Justice Waite's statement that "[t]he court does not wish to hear argument on the question whether the [equal protection clause] applies to these corporations. We are all of the opinion that it does" (Santa Clara 1886: 396). Some scholars identified this as an application of the real entity view to corporations, but Professor Horwitz has shown by examining Justice Field's opinion in the court, below, that it actually represented an application of the aggregate view (Horwitz 1985: 176–83). Specifically, Field held that the equal protection clause must apply to corporations for the following reasons:

> Private corporations consist of an association of individuals united for some lawful purpose, and permitted to use a common name in their business and have succession of membership without dissolution. . . . But the members do not, because of such association, lose their right to protection, and equality of protection. . . . So, therefore, whenever a provision of the constitution or of a law guarantees to persons protection in their property . . . the benefits of the provision are extended to corporations; not to the name under which different persons are united, but to the individuals composing the union. The courts will always look through the name to see and protect those whom the name represents.
>
> (Santa Clara 1883: 402–3, citing Deveaux)

A clearer statement of the aggregate view can hardly be imagined; most remarkable is Field's reliance on Deveaux despite the fact that the Supreme

Court overturned its results forty years earlier. Similarly, in Pembina Consolidated Silver Mining & Milling Co. v. Pennsylvania, decided two years later in 1988, Justice Field, for the Court, stated that "[u]nder the designation of person there is no doubt that a private corporation is included. Such corporations are merely associations of individuals united for a special purpose . . ." (Pembina: 189).[20]

The artificial entity view, however, was also raised in these cases. In Santa Clara, the railroad corporations made the argument that because they were operating under special congressional legislation they should be regarded as an extension of the federal government and therefore California could not tax them (Santa Clara 1886: 387). Field rejected this view, citing Trustees of Dartmouth College, but noted that "when the instrumentality is the creation of the state, a corporation formed under its laws, and is employed or adopted by the general government for its convenience . . . it remains subject to the taxing power of the state" (Santa Clara 1886: 389). And notably, in Pembina, Field followed Taney in rejecting the argument that the privileges and immunities clause applied to corporations because they were not "citizens," even though the aggregate view he adopted in Santa Clara might have led to the contrary position. Instead, Field emphasized the relationship between the corporation and the incorporating state under the artificial entity view:

> [T]he term citizens, as used in this clause, applies only to natural persons, members of the body politic owing allegiance to the State, not to artificial persons created by the legislature, . . . a grant of corporate existence was a grant of special privileges to the corporators, enabling them to act for certain specified purposes as a single individual, and exempting them, unless otherwise provided, from individual liability.
>
> (Pembina: 187–8)

All three views of the corporation appear in Hale v. Henkel, decided by the Supreme Court in 1906. The issue was whether an agent of a corporation could invoke the Fifth Amendment privilege against self-incrimination or the Fourth Amendment protection against unreasonable search and seizure in the name of the corporation. On the Fifth Amendment issue, the Court held that the right against self-incrimination does not apply to corporations:

> The right of a person under the Fifth Amendment to refuse to incriminate himself is purely a personal privilege of the witness. . . . The question whether a corporation is a "person" within the meaning of this Amendment really does not arise . . . since it can only be heard by oral evidence in the person of some one of its agents or employees.
>
> (Hale: 69–70)

This is closest to the real entity view because it rejects (like Marshall in Dandridge) the aggregate position of looking through a corporation to its shareholders and

takes into account the special characteristics of the corporation itself. On the other hand, on the Fourth Amendment issue, the Court at first emphasized the artificial entity view, using it to justify regulation by the state:

> Conceding that the witness was an officer of the corporation under investigation, and that he was entitled to assert the rights of the corporation with respect to the production of its books and papers, we are of the opinion that there is a clear distinction in this particular between an individual and a corporation, and that the latter has no right to refuse to submit its books and papers for an examination at the suit of the State. The individual may stand upon his constitutional rights as a citizen. . . . Upon the other hand, the corporation is a creature of the State. It is presumed to be incorporated for the benefit of the public. It receives certain special privileges and franchises, and holds them subject to the laws of the State and the limitations of its charter. Its powers are limited by law. It can make no contract not authorized by its charter. Its rights to act as a corporation are only preserved to it so long as it obeys the laws of its creation. There is a reserved right in the legislature to investigate its contracts and find out whether it has exceeded its powers. It would be a strange anomaly to hold that a State, having chartered a corporation to make use of certain franchises, could not in the exercise of its sovereignty inquire how these franchises had been employed, and whether they had been abused, and demand the production of the corporate books and papers for that purpose. . . . While an individual may lawfully refuse to answer incriminating questions unless protected by an immunity statute, it does not follow that a corporation, vested with special privileges and franchises, may refuse to show its hand when charged with an abuse of such privileges.
>
> (Hale: 74–5)

Having clearly stated its reasons for limiting the application of the constitutional right, however, the Court suddenly reverts to the aggregate view when facing the question whether corporations have any Fourth Amendment rights at all:

> [W]e do not wish to be understood as holding that a corporation is not entitled to immunity, under the Fourth Amendment, against unreasonable searches and seizures. A corporation is, after all, but an association of individuals under an assumed name and with a distinct legal entity. In organizing itself as a collective body it waives no constitutional immunities appropriate to such body. Its property cannot be taken without compensation. It can only be proceeded against by due process of law, and is protected, under the Fourteenth Amendment, against unlawful discrimination. Corporations are a necessary feature of modern business activity, and their aggregated capital has become the source of nearly all great enterprises.
>
> (Hale: 76; citations and emphasis omitted)

What can explain this remarkable oscillation between the three views? The key is in the last sentence quoted. As noted above, the period between 1890 and 1906 marked the height of the debate on the rise of the great corporations. The Court is trying to strike a balance between the rights of the corporations, which can best be protected under either the aggregate or the real entity views, and the regulatory power of the state, which is best reflected in the artificial entity view. On the one hand, as the Court states, "Corporations are a necessary feature of modern business activity" and must be protected (Hale: 76). On the other hand, the right of the state to regulate must also be preserved, especially since the context of Hale v. Henkel was an antitrust investigation into two major corporations, the American Tobacco Company and McAndrews & Forbes Inc.

Ultimately, however, the real entity view prevailed.[21] This involved first the rejection of the aggregate view. For example, in Southern Railway Co. v. Greene, decided in 1909, the Court came out clearly for the position that the corporation as such was entitled to constitutional protection under the equal protection clause, without any reference to its shareholders: "the corporation . . . is within the meaning of the Fourteenth Amendment, a person within the jurisdiction of the state of Alabama, and entitled to be protected against any statute of the State which deprives it of the equal protection of the laws" (Southern Railway: 417).[22]

Once again, the triumph of the real entity view can be explained by several factors. The aggregate view was raised by Field and others to protect the rights of corporations, but it was even more incongruous in the context of the mega-corporations of the 1890s, with thousands more shareholders than in the pre-Civil War days. It also gave the corporation too many rights *vis-à-vis* the state, as seen in Hale v. Henkel. The artificial entity view gave the state too much power to regulate corporations, as the Hale v. Henkel Court came to realize when it laid out its implications. The real entity view was the most congruent with business realities as well as the one most suited to some balance between corporations and the state. By 1909, it was well established as the dominant view of the corporation, as reflected in contemporary debates surrounding the enactment of the corporate tax (Avi-Yonah 2004: 1215–25).

The rise of the real entity view is also reflected in two other contemporary developments: the rise of the business judgment rule, and the decline of the *ultra vires* doctrine.[23] The business judgment rule rejected the aggregate view in holding that the board of directors held powers that were not delegated from the shareholders and that shareholders could not normally call into question the exercise of those powers. The *ultra vires* doctrine represented the ability of the state to require corporations to adhere to their charter, and was thus based on the artificial entity view; its decline thus reinforced the rejection of that view.

Under the business judgment rule, courts will not second guess good-faith decisions made by independent and disinterested directors. To earn the protection of the business judgment rule, directors must prove that they are financially disinterested, were duly informed, they acted "in good faith," and finally that the directors acted "in the honest belief that the action taken was in the best interests of the company."[24]

The first full statement of the business judgment rule was made in Leslie v. Lorillard, decided by the New York Court of Appeals in 1888. This statement was expanded a year later in Beveridge v. N.Y. Elevated R.R, and the rule became well established, so that by 1905 the court in Siegman v. Elec. Vehicle Co. wrote that "it is [the board's] judgment, and not that of its stockholders outside of the board of directors . . . that is to shape [a corporation's] policies or decide upon its corporate acts. This principle is not disputed, and the citation of authorities in its support is unnecessary."[25] The rule reflected the real entity view, which equates the corporation with its management, and rejected the aggregate view of the corporation as an aggregate of its shareholders.

The one potential limitation on the power of the board was the *ultra vires* doctrine, which held that a board could not act contrary to the powers conferred on it by the state. The *ultra vires* doctrine thus represented the artificial entity view. The doctrine originated in the pre-Civil War era,[26] but became prominent in the arguments on the relationship of the state and the corporation in the 1880s and 1890s (Cook 1887: chs 19 and 38). The artificial entity argument for upholding the limitation was stated clearly by the New York Court of Appeals in 1888:

> Corporations are great engines for the promotion of the public convenience, and for the development of public wealth, and, so long as they are conducted for the purposes for which organized, they are a public benefit; but if allowed to engage, without supervision, in subjects of enterprise foreign to their charters, or if permitted unrestrainedly to control and monopolize the avenues to that industry in which they are engaged, they become a public menace, against which public policy and statutes design protection.
>
> (Leslie: 531–3)

The Supreme Court upheld the doctrine in the following year on the ground of corporate abuses:

> . . . while valuable services have been rendered to the public by this class of organizations, which have stimulated their formation by numerous special acts, it came at last to be perceived that they were attended by many evils in their operation as well as much good, and that the hasty manner in which they were created by the legislatures, sometimes with exclusive privileges, often without due consideration and under the influence of improper motives, frequently led to bad results.
>
> (Oregon Railway: 20–1)

The reference to corporate abuses relates to the rise of trusts, and indeed the *ultra vires* doctrine was used to dissolve sugar and oil trusts under New York and Ohio law (Dwight 1888; Cook 1893).[27] In 1895, however, the Supreme Court rejected an antitrust challenge to the sugar trust on the grounds that the Sherman Act applied only to corporations engaged directly in interstate commerce (United

States v. E.C. Knight Co.: 17). In 1896, the Court rejected an *ultra vires* challenge to the ability of the Union Pacific Railway to lease its tracks for 999 years to another railroad, when the charter would not permit an outright sale (Union Pac. Ry. Co. v. Chicago, Rock Island & Pac. R.R.: 594–6). This literal decision significantly reduced the power of the *ultra vires* doctrine (Horwitz 1985: 186–8).[28]

The ultimate demise of the doctrine resulted not from a court decision but from the competition among states to attract corporate charters, which was begun by New Jersey in 1890 and continued by Delaware in the 1900s (Larcom 1937: 1–26; Steffens 1906: 209; Keasbey 1899: 209–11).[29] This competition meant that New Jersey and Delaware had every incentive to relax any limiting elements in their charters that restricted the power of corporate management.[30] Thus, for example, the long-lasting prohibition against corporations owning stock in other corporations, which led to the necessity of "trusts," was eliminated by New Jersey in its 1890 law (Cook 1898).[31] As a result, although the Supreme Court still held in De La Vergne Refrigerating Mach. Co. v. German Sav. Inst. (1899) that such a combination was *ultra vires* under New York law, this holding became rather meaningless since most large, publicly traded corporations were incorporated in New Jersey. As the New Jersey statute explains:

> It was formerly the rule in this State that acts of a corporation in excess of its express powers, or those necessarily implied, were void, and contracts which were *ultra vires* the corporation were incapable of enforcement or ratification. . . . This rule no longer obtains.
>
> (General Corporation Law of New Jersey sec. 2 (1896))

The decline of the *ultra vires* doctrine was sealed by the spread of corporate laws permitting incorporation "for any lawful purpose." With the doctrine gone, the artificial entity view of the corporation became less plausible, and the real entity view reigned supreme again (Machen 1911: 257–8).[32]

### *A failed transformation: the hostile takeover crisis*

In 1926 John Dewey published an article in the Yale Law Journal in which he dismissed as irrelevant the debate among the aggregate, artificial entity, and real entity views of the corporation. These views, he explained, could be deployed to suit any purpose; and he used examples relying on the cyclical nature of these theories. His conclusion was that theory should be abandoned for an examination of reality (Dewey 1926: 673).

Dewey was influential in that the theoretical debate on corporate personality largely disappeared until the 1970s. As a practical matter, however, the real entity view predominated for large, publicly traded corporations. The board ran the corporation as it saw fit, protected from the shareholders by the separation of ownership from management noted by Berle and Means in the 1930s, and by the business judgment rule, and protected from the state by the relaxation of corporate law limits begun by New Jersey and continued by Delaware.[33]

The next significant practical change in this state of affairs only arose in the 1980s, when corporate America witnessed a takeover wave that included large components of hostile takeovers, bust-ups, and leveraged buyouts. During that era, it suddenly became possible for hostile raiders to threaten takeovers of even the largest corporations. After RJR Nabisco was taken private for $25 billion in 1988, it was clear that no board was safe. Legislatures of many states have attempted to protect corporations in their local states from hostile takeovers and have passed anti-takeover or constituency statutes. The typical constituency statute provides that in acting in the best interests of the corporation, the directors may take into account the interests of a variety of constituencies other than shareholders, including employees, the communities in which the corporations are located, customers, and suppliers (Orts 1992; Hansen 1991). The diminished board's power and the anti-takeover legislation commenced once again the debates on the nature of the corporation and its relationship to the shareholders and the state. Once again all three theories of the corporation reappeared, as can be seen if one examines the following seminal cases decided between 1982 and 1989 by the Supreme Courts of the United States and of Delaware.

Edgar v. MITE Corp., decided by the U.S. Supreme Court in 1982, involved the constitutionality of an anti-takeover act enacted by the state of Illinois. Under the Illinois Business Take-Over Act, a hostile tender offer for the shares of a company covered by the act had to be registered by the Secretary of State and the offeror had to give both the target and the state a twenty-day notice during which only the target could communicate with its shareholders regarding the offer. The act applied both to corporations whose 10 percent of their shareholders were residents of Illinois and to corporations that were either incorporated in the state or had their principal office in it. The MITE corporation made a hostile offer for an Illinois corporation and refused to comply with the act, arguing that it violated the commerce clause.

The Supreme Court agreed with MITE. Writing for a 5–4 majority, Justice White held that the Illinois act was unconstitutional because it could apply to tender offers that did not affect a single Illinois shareholder; "the state has no legitimate interest in protecting nonresident shareholders" (Edgar: 644). Moreover, the fact that the target corporation was an Illinois corporation was irrelevant since state regulation only applied to the corporation's "internal affairs": "[t]ender offers contemplate transfers of stock by stockholders to a third party and do not themselves implicate the internal affairs of the target company" (Edgar: 645). Instead, the focus should be entirely on the impact of blocking the tender offer on the company's shareholders and their relationship with management:

> The effects of allowing the Illinois Secretary of State to block a nationwide tender offer are substantial. Shareholders are deprived of the opportunity to sell their shares at a premium. The reallocation of economic resources to their highest valued use, a process which can improve efficiency and competition,

> is hindered. The incentive the tender offer mechanism provides incumbent management to perform well so that stock prices remain high is reduced.
>
> (Edgar: 643 (referring to Easterbrook and Fischel 1981: 1173–4; and Fischel 1978: 5, 27–8))

This part of the opinion clearly reflects the aggregate view: the focus is entirely on the impact on the corporation's shareholders, and the corporation itself (including its management) barely exists. The market for corporate control is praised because of its ability to overcome the agency cost problem and the incentive it provides for management to maximize stock prices. Moreover, White quotes the work of Easterbrook and Fischel, who are among the principal proponents of the "nexus of contracts" theory of the corporation, according to which the corporation is merely a convenient legal term for a series of contracts, the most important of which is the contract between shareholders and management (Fischel 1982: 273).[34]

In addition, this part of the opinion, which rejects both the artificial entity and the real entity theories, evoked some misgivings on the part of Justice Powell, even though he joined it to provide the crucial fifth vote. Powell noted that in some cases the state may have a legitimate interest because the corporation has a real presence that goes beyond a contract between management and the shareholders, reflecting both the artificial and real entity views:

> I join Part V-B because its Commerce Clause reasoning leaves some room for state regulation of tender offers. This period in our history is marked by conglomerate corporate formations essentially unrestricted by the antitrust laws. Often the offeror possesses resources, in terms of professional personnel experienced in takeovers as well as of capital, that vastly exceed those of the takeover target. This disparity in resources may seriously disadvantage a relatively small or regional target corporation. Inevitably there are certain adverse consequences in terms of general public interest when corporate headquarters are moved away from a city and State.
>
> (Edgar: 646)

As a footnote, Powell added:

> The corporate headquarters of the great national and multinational corporations tend to be located in the large cities of a few States. When corporate headquarters are transferred out of a city and State into one of these metropolitan centers, the State and locality from which the transfer is made inevitably suffer significantly. Management personnel – many of whom have provided community leadership – may move to the new corporate headquarters. Contributions to cultural, charitable, and educational life – both in terms of leadership and financial support – also tend to diminish when there is a move of corporate headquarters.
>
> (Edgar: 646)

The artificial entity view of the corporation was emphasized five years later, when Powell had the opportunity to translate these misgivings into an opinion for the Court. CTS Corp. v. Dynamics Corp. of America (1987) involved a so-called "second generation" anti-takeover statute, i.e., one that was drafted to get around the problems which the Illinois statute struck down in Edgar v. MITE. The Indiana statute applied only to corporations incorporated in Indiana, which have a specified number of shareholders within the state, and which opt for its protection. Under the statute, an acquirer who acquired "control shares" in such an Indiana target could vote them only with the approval of a majority of the pre-existing disinterested shareholders, to be obtained in a meeting within fifty days after the acquisition.

The court of appeals followed Edgar and declared the statute unconstitutional under the commerce clause, because it interfered with the market for corporate control (Dynamics Corp. of Am. v. CTS Corp. 1986: 264). The Supreme Court reversed. Justice Powell, writing for a 5–4 majority, stated:

> No principle of corporation law and practice is more firmly established than a State's authority to regulate domestic corporations, including the authority to define the voting rights of shareholders. . . . We think the Court of Appeals failed to appreciate the significance for Commerce Clause analysis of the fact that state regulation of corporate governance is regulation of entities whose very existence and attributes are a product of state law. As Chief Justice Marshall explained: "A corporation is an artificial being, invisible, intangible, and existing only in contemplation of law. Being the mere creature of law, it possesses only those properties which the charter of its creation confers upon it, either expressly, or as incidental to its very existence. These are such as are supposed best calculated to effect the object for which it was created."
>
> (CTS Corp. 1987: 89 (citing Dartmouth College))

Powell thus rejected the view that states do not have the right to regulate transactions affecting shareholders, including shareholders of other states. He argued that the "free market system depends at its core upon the fact that a corporation . . . is organized under, and governed by, the law of a single jurisdiction. . . . A State has an interest in promoting stable relationships among parties involved in the corporations it charters" (CTS Corp. 1987: 90–91). He explicitly rejected the market for corporate control and its underlying aggregate theory:

> The Constitution does not require the States to subscribe to any particular economic theory . . . there is no reason to assume that the type of conglomerate corporation that may result from repetitive takeovers necessarily will result in more effective management or otherwise be beneficial to shareholders . . . the very commodity that is traded in the "market for corporate control" – the corporation – is one that owes its existence and attributes to state law.
>
> (CTS Corp. 1987: 92–4 (citations and emphasis omitted))

This entire opinion, with its quotation from Dartmouth College, is clearly based on the artificial entity view that the corporation owes its existence to the incorporating state and that the state may therefore regulate it, including in ways that affect shareholders' ability to sell their shares. Not surprisingly, Justice White dissented, arguing that while the statute may help Indiana corporations "particularly in helping those corporations maintain the status quo," it is inimical to the interests of the shareholders and constitutes "economic protectionism" (CTS Corp. 1987: 98–100).

At the same time, the battle for corporate control appeared also in state law, and the most important state in this regard was Delaware, in which most major U.S. corporations are incorporated. In Unocal Corp. v. Mesa Petroleum Co., Delaware Supreme Court held in 1985 that a board which wishes to resist a hostile takeover bid can enjoy the protection of the business judgment rule if it can prove that the defensive tactic adopted by the board was reasonable and in relation to the level of the posed threat. The Delaware Supreme Court also made clear, however, that in deciding whether there is a threat to the corporate entity, the board is invited to consider "the impact on ... creditors, customers, employees, and perhaps even the community generally" (Unocal: 955). In other words, the Delaware Supreme Court rejected the aggregate view (by rejecting the shareholders' primacy doctrine) in favor of the real entity view, in which management is free to consider the interests of the "corporation" that include also the interests of non-shareholder constituencies. This view received its final affirmation in 1989, when the Supreme Court of Delaware issued an opinion in Paramount Communications, Inc. v. Time, Inc. that in practice ended the hostile takeover boom. Paramount had made a $175 (later raised to $200) per share offer for Time at the time when Time was about to enter into a $70 per share merger with Warner. Paramount argued that, under the previous decisions of the Delaware Supreme Court in Unocal (1985) and Revlon (1986), Time was "up for sale" and therefore the business judgment rule was suspended and Time's board was required to maximize shareholder value by accepting the much higher Paramount bid.

The Delaware Supreme Court held in favor of Time, allowing Time's board to reject the lucrative hostile offer from Paramount and to pursue other non-shareholder interests. It stated that:

> Two key predicates underpin our analysis. First, Delaware law imposes on a board of directors the duty to manage the business and affairs of the corporation. This broad mandate includes a conferred authority to set a corporate course of action, including time frame, designed to enhance corporate profitability. Thus, the question of "long-term" versus "short-term" values is largely irrelevant because directors, generally, are obliged to charter a course for a corporation which is in its best interest without regard to a fixed investment horizon. Second, absent a limited set of circumstances as defined under Revlon, a board of directors, while always required to act in an informed manner, is not under any per se duty to

> maximize shareholder value in the short term, even in the context of a takeover.
>
> (Paramount: 1150 (citations omitted))

The court thus rejected the view that maximizing short-term shareholder value was always required; instead, the board was permitted to pursue its view of the best long-term corporate strategy (Paramount: 1154). Thus, the board was permitted to prefer preservation of the "Time culture" (its stated goal, instead of maximizing the cash return to shareholders). This effectively killed the takeover threat, because any board could find good long-term share value maximization reasons to reject a superior cash bid. The Delaware court, in enhancing managerial power, effectively endorsed the real entity view: a corporation was an entity with its own corporate culture, which should not be subordinated to the shareholders or to the state.

This view was ratified when the ALI Corporate Governance Project (1994) adopted a rule that corporate boards may take into account the interests of other "stakeholders," not just the shareholders (§ 2.01(b)(3)).

Why did the real entity view prevail? A plausible answer was that corporate management determines the state of incorporation, and therefore the Delaware Supreme Court felt that it had to side with management once the U.S. Supreme Court had approved the anti-takeover laws of other states, lest corporations choose to relocate there. It seems unlikely, however, that this was the only reason; Delaware is very well established as the preferred state of incorporation, and stock values would likely decrease if shareholders perceived that managements were leaving Delaware just to protect themselves. Instead, it seems likely that the Delaware Supreme Court genuinely believed that a corporation like Time had a corporate existence and culture with implications for other stakeholders, and therefore rejected the aggregate view equating the corporation with its shareholders. In that way, its concerns were similar to those raised by Justice Powell in his concurrence in Edgar: a corporation is more than a "nexus of contracts," and courts and legislatures are allowed to take the interests of other stakeholders into account (Edgar: 646–7).

## *Fourth transformation: from national corporations to multinational enterprises*

The last transformation in the nature of the corporation began in the 1950s and is still going on, so that its ultimate outcome is hard to judge. This is the transformation from corporations based mostly in one country to multinational enterprises based in many countries.

Multinationals, in the sense of corporations owning assets overseas, have existed since the seventeenth century (Harris 2000: 39–59). As recently as the 1950s, however, the shareholders and other sources of capital, the management, most of the production facilities, and most of the markets of even large multinationals tended to be in one country, so that "what was good for G.M. was

good for America" (Avi-Yonah 2002: 1796). By the 1990s, however, this had changed profoundly. As more countries opened up to foreign direct investment, communications improved, many products became lighter and easier to ship, and more and more corporations became "globalized." In a globalized multinational corporation, the sources of capital are in many countries: the shares of large multinationals trade on as many as twenty exchanges, and borrowing facilities are similarly diversified. Research and development and production facilities are likewise spread throughout the globe, as are markets. The only thing that usually ties a modern multinational to its home country is the location of management.

In this context, the debate over the nature of the corporation has re-opened. There is abundant academic writing on the relationship between multinationals and the state, and most writers from both the left and the right concede that this relationship has changed profoundly so that the home state (the state of incorporation) has become powerless to control "its" multinationals; it is hard even to identify to which country multinationals "belong" (Graham and Krugman 1995: 86–93; Reich 1990: 53–4; but see Tyson 1991: 47–9). On a practical level this situation has led to attempts by home states to control the behavior of multinationals abroad in areas as diverse as trading with the enemy, antitrust, corruption, and others, with varying success (Muchlinski 1996: 126–56; Vernon 1998: 215–16; Avi-Yonah 2003: 17–31; and Blumberg 1993: 193–4). The most recent development in this regard has been "inversion" transactions, in which the management changes the country of incorporation of a multinational's parent corporation. These transactions are undertaken primarily for tax reasons, but they have corporate governance implications as well (Avi-Yonah 2002: 1793–7). Specifically, the artificial entity theory becomes hard to maintain when management can pick weak countries like Bermuda as the country of incorporation for the parent of a multinational.

The relationship with shareholders has also undergone changes as shareholders now tend to come from many countries. One implication of this has been that the securities laws of the weakest country tend to dominate because of cross-country price arbitrage (Licht 1998: 591). Another is academic proposals to let management choose the country of securities law as well as the country of incorporation (Fox 1999: 1337). On a practical level globalization has led the SEC to relax requirements for some foreign issuers (Fox 1998: 713–14). This trend has tended to weaken the applicability of the aggregate view as well. It is hard to predict where these trends will lead, but at the moment they appear once more to favor the real entity view.

To summarize: throughout all the transformations we have studied, the same pattern recurs. As the relationship of the corporation to the state, to society and to its members or shareholders changes, all three views of the corporation emerge, submerge and then re-emerge in a slightly different but fundamentally similar form. In the end, however, the real entity view prevails.

Why does the real entity view prevail? First, the real entity view prevails because it shields corporate management from undue interference from both

shareholders and the state, and thus it reflects the power of corporate management. One way of looking at the transformations outlined above is that both the artificial entity and aggregate views were advanced in order to limit the power of management. The artificial entity view was usually brought forward in order to enable the state to regulate corporations, and the aggregate view was usually advanced to enhance the power of shareholders, although sometime it was used to give corporations rights that normally belong only to individuals. The ultimate success of the real entity view resulted from the fact that it gave more power to management than the other views. Corporate management wields political power and it influences the outcome of the debate; judges again and again refer to the importance of corporations, by which they mean corporate management. But the very fact that corporate management wields this power shows that there is another reason why the real entity view prevails: it represents a better approximation of reality than the artificial entity view and aggregate view, and this view became a better approximation as the corporate form evolved. Roman or medieval corporations could plausibly be seen as merely creatures of the state because of the state's role in creating them, or they could be seen as mere aggregations of their members because the members also managed the corporation. These views are much less plausible today, however, since the state plays only a minimal role in creating corporations and that role is sharply constrained by management's ability to shift the location of incorporation. The shareholders, meanwhile, are often spread all over the globe and are clearly separate from the corporate entity. The real entity view prevailed because it was more real than the others.[35] This observation enables us to move from the historical to the normative part of the discussion and ask what implications the reality of corporations has for corporate law and regulation.

## Normative implications for corporate social responsibility (CSR)

A page of history may be worth a volume of logic[36] as far as explanatory power is concerned, but Justice Holmes also conceded that history *per se* has no normative power (Holmes 1897: 459). What are the implications of these cyclical transformations of corporate theory for the problem of CSR? Can we draw any conclusions on the legitimacy of CSR from the history described above? It could be argued that the answer is "yes," and that the dominance of the real entity theory for most of corporate history has far-reaching implications for the legitimacy of CSR activities for the reasons explained below.

### *The three theories and CSR*

Each of the three theories of the corporation permits a different level of CSR, as indicated in Table 10.1.

The first type of CSR involves activities that can clearly and demonstrably benefit shareholders in the long run. For example, actions that prevent environmental

*Table 10.1* Theories of the corporation and corporate social responsibility (CSR)

| *Theory* | *Aggregate* | *Artificial* | *Real* |
|---|---|---|---|
| For long-run benefit of shareholders | Yes | Yes | Yes |
| Not for shareholders; Corporation responsible | No | Yes | Yes |
| Not for shareholders; Corporation not responsible | No | No | Yes |

disasters or comply with legal and ethical rules can have a significant positive effect in preventing disastrous corporate calamities, even if they cost money in the short run. Thus, even proponents of the aggregate theory, the currently dominant theory of the corporation in academic circles, would support this type of CSR.

The second type of CSR involves activities that are designed to mitigate social harms the corporation was responsible for, even when there is no direct legal responsibility, and when no benefit to the shareholders can be shown. Under the aggregate theory, such activities should not be permitted because they do not benefit shareholders. But under the artificial entity theory, since it emphasizes the benefits of corporate existence derived from the state, an implicit contract can be inferred that the corporation will help the state in mitigating harms that it causes even in the absence of legal responsibility. Otherwise, the state will have to bear this burden imposed by the corporation it created.

Finally, the third type of CSR involves activities like AIDS prevention, for which the corporation is not responsible and which in most cases do not benefit its shareholders, even in the long run. This type of CSR would not be permitted under the aggregate or artificial entity theories. But under the real entity theory, since the corporation is regarded as one person in law just like individuals,[37] it is permitted to act philanthropically just like individuals are, and should, in fact, be praised to the extent it does so (White 1999: 111–23).

Thus, under the real entity theory, even CSR activities that have nothing to do with benefiting shareholders or with direct corporate responsibility are permitted. This still requires an answer to the two arguments advanced by Levitt, Friedman, Jensen, and their colleagues.[38]

### *Real entity as a response to CSR's opponents*

First, there is the argument that the money being spent on CSR belongs to the shareholders and therefore management have no right to spend it according to their preferences in ways that are not related to maximizing shareholder value. However, as long as the corporation's CSR activities are adequately disclosed to the shareholders (and the securities laws are designed to assure such disclosure is made),[39] it is not clear that they have a right to complain. If the shareholders do not like the firm engaging in CSR activities, they can sell the shares and invest solely in firms that do not engage in such activities. Easterbrook and Fischel phrased it adequately:

> What is the goal of the corporation? Is it profit, and for whom? . . . Our response to such questions is: who cares? If the New York Times is formed to publish a newspaper first and make a profit second, no one should be allowed to object. Those who came in at the beginning consented, and those who came later bought stock the price of which reflected the corporation's tempered commitment to a profit objective . . . Similarly, if a bank is formed with a declared purpose of giving priority to loans to minority-owned businesses or third-world nations, that is a matter for the ventures to settle among themselves.
>
> (Easterbrook and Fischel 1991: 35–6)

Moreover, it can be argued that the majority of current shareholders, namely those who invest through mutual funds and pension funds, invest primarily to obtain a secure return and not for maximum, but risky, gains. In this sense, most shareholders today are more like bondholders or preferred shareholders, who care more about a stable return than about value maximization. For those shareholders, firms that promise a secure, reasonably high return are a good investment, even if they reduce the chances of obtaining returns over that limit by engaging in CSR. Those shareholders who seek to maximize returns are then free to invest in firms that do not engage in CSR.

Second, there is the argument that if firms are free to engage in CSR, it will be more difficult to evaluate management performance since there will be no single benchmark like earnings per share. This may be true, but in a complex world, we are used to evaluating leaders on more than one benchmark. We would not seriously argue that political leaders, for example, must be evaluated only on their economic performance and on no other measure. If we can use complex measures to evaluate politicians, we can do the same for CEOs.

In addition, as Chen and Hanson point out, there is an internal inconsistency in Milton Friedman's argument, because if markets are efficient they should prevent managers from engaging in actions that are not in the best interest of shareholders (Chen and Hanson 2004: 42–66). Friedman may, in fact, have believed that to be the case, but the dominance of the real entity view of the corporation through 2,000 years of corporate history suggests that management usually finds a way to do as they wish, including engaging in CSR when it may not be in the long-term interest of shareholders. Since the courts are unable to effectively police such behavior and markets are an insufficient constraint, it is unclear what in practice is gained by arguing in favor of shareholder primacy and against CSR.

Thus, if the historical argument advanced above is correct, and real entity theory is, in fact, the dominant theory of the corporation for most of its history because it is a more accurate description of reality than either the artificial entity or aggregate theories, this can justify CSR to a much greater extent than is commonly accepted by most corporate law academics. Why, then, has the aggregate (nexus of contracts) theory achieved such success in U.S. academic circles? The scholars Blair and Stout provide an appealing answer that emphasizes the real entity theory and management power:

> We do not think it is an accident that the idea of shareholder primacy has become increasingly popular among academics during this period. Our theory suggests that the shift in the balance of power in boardrooms toward shareholders is the result not of directors' sudden recognition that shareholders are in fact "owners" of the corporation, however, but of changing economic and political forces that have improved shareholders' relative bargaining power vis-a-vis other coalition members. . . . Thus, at a normative level our story cautions against attempts to "reform" corporate law either by contractarians who want to enhance shareholders' power over directors, or progressives who want to give other stakeholders greater control rights. Strikingly, corporate law itself has proven remarkably immune to both sorts of proposals, and continues to preserve directors' discretion to act as mediators among all relevant corporate constituents (Blair and Stout 1999: 327–8).

Finally, the purpose of this chapter has been to show that, even in the U.S. context, and despite the current dominant legal theory of nexus of contracts, the aggregate theory has not always been dominant. In fact, throughout most of the history described above, the real entity theory was the dominant one, and it can be argued that in practice most corporations are still operating on the basis of the real identity theory, not the aggregate one. Thus, CSR is most easy to justify in all its forms on the basis of the real identity theory of the corporation and is likely to remain practiced for the future. The debate on CSR should, therefore, shift from whether CSR is acceptable to how to make it more accountable and effective in obtaining social goals – but that is an issue for another day (Walsh and Avi-Yonah 2005).

## Notes

* This chapter is based in part on Reuven S. Avi-Yonah (2005), "The Cyclical Transformations of the Corporate Form: A Historical Perspective on Corporate Social Responsibility," 30, *Del. J. Corp. L.* 767.

1 United Nations (2001).

2 Berle 1931: 1049 (stating directors' powers are to be held in trust for the shareholders as the only beneficiaries of the corporation); Dodd 1932: 1145 (opposing shareholders primacy and arguing that corporations responsible for many constituencies); and Berle 1932: 1365 (stating that imposing strict fiduciary duty only toward the shareholders is the sole way to prevent directors from oppressing shareholders). But see Berle 1954 at note 33 hereinafter (conceding to Dodd's contention).

3 Mitchell 1995; Branson 2002a: 1217–22; Branson 2002b: 122; Williams 1999: 1267–8.

4 The classic case affirming this "shareholder primacy" doctrine is Dodge v. Ford Motor Co. (1919). See also the classic debate between Berle and Dodd, *supra* note 2. See, e.g., Easterbrook and Fischel 1991: 15–22 (stating that shareholders, as residual claimants, have implicitly contracted for promise that firm will maximize profits in long run); Manne and Wallich 1972 (noting that social responsibility of corporations is shareholder wealth maximizing); Black and Kraakman 1996: 1921 (arguing that principal goal of corporate law is to maximize shareholder wealth); Hansmann and Kraakman 2001: 441, 449–51 (stating that shareholder primacy is likely to dominate

future development of corporate law); Greenwood 1996: 1023 ("[A]ll but the communitarians agree that virtually the sole task of corporate law is to ensure that managers act as agents for the shareholder owners."); Romano 1987: 113 (asserting that the core goal of corporate law is to maximize equity share prices); Bebchuk and Fried 2004: 201–16 (discussing the need to align managerial incentives with shareholder interests). For arguments on the other side, Blair and Stout 1999: 250–1, 253–4, 288, 310–11, 320 (arguing that a corporation should be envisioned as a team production, in which the board acts as a "mediating hierarchy" mediating among the various competing constituencies of the corporation, and asserting that the shareholders' primacy norm is an error of legal analysis and that in fact corporate law requires directors to serve not only the shareholders' interests, but also those of other stakeholders); Millon 1990: 261–2 (praising case law that reaffirms directors' discretion to consider nonshareholder interests); Mitchell 1992: 630–43 (arguing that courts should modify corporate law to grant stakeholders standing to sue directors when the former are harmed by corporate action); O'Connor 1993: 936–65 (arguing that corporate law should be changed to encourage employee representation on the board and standing to sue); Williams 2002 (stating that one of the striking features of American corporate law is how little real control shareholders have, given that they are the "owners" of the corporation).

5 See also The Business Roundtable, Statement on Corporate Governance and American Competitiveness, 46 *Bus. Law* 241, 244 (1990) (recognizing that corporations serve the interests of both its shareholders and society).

6 A similar shift also occurred during the nineteenth century in Europe.

7 Limited liability generally was embraced by the civil law earlier than in the Anglo-American world. The Napoleonic *Code de Commerce* in 1807 was the first enactment in the Western world that provided for limited liability for stock corporations such as the *sociétés anonyms*. See C. Com. art. 33 (1807) (France). Subsequently, limited liability was quickly adopted throughout Europe. See Blumberg 1986: 596.

8 To be sure, notwithstanding the significant distinction between corporation and enterprise (see Berle 1947, reprinted in this volume), for the purposes of this paper the term corporation includes enterprise, and in many respects the analysis of the real entity of the corporation applies also to the legal entity of the enterprise.

9 For different perspectives on the development of the corporate form, see Schultz 1951: 88–102; Gierke 1977: 128–31.

10 For detailed discussion on the pre-nineteenth-century period see Avi-Yonah 2005: 773–82.

11 Tipling v. Pexall, 3 *Bulstrode* 233 (1614) ("the King creates them"). For an example of a charter enumerating corporate legal rights, see Case of Sutton's Hospital, 77 Eng. Rep. 937 (1612) (KB).

12 Limited liability was well established by Blackstone's time for royally chartered corporations. "The debts of a corporation, either to or from it, are totally extinguished by its dissolution, so that the members cannot recover, or be charged with them, in their natural capacities" (Blackstone 1765: 462).

13 See for example Angell and Ames 1832, 35: "In no country have corporations been multiplied to so great an extent, as in our own. . . . There is scarcely an individual of respectable character in our community, who is not a member of, at least, one private company or society which is incorporated. . . . Acts of incorporation are moreover continually solicited at every session of the legislature."

14 This was subject to one limitation, the "trust fund" doctrine, which held that the capital stock of a corporation was to be held in trust for paying corporate debts and thus could not be distributed to shareholders while debts were outstanding. See Wood v. Dummer.

15 See also Marshall v. Baltimore and Ohio R.R (1853) at 329 (holding that for diversity purposes a corporation should be deemed a resident of its place of incorporation).

This led to the current rule, adopted in 1958, under which a corporation is for diversity purposes a citizen of both the state it is incorporated in and the state in which it has its principal place of business. 28 U.S.C. § 1332(c)(1) (2001).

16 See also Chief Justice Shaw's statement in Burrill v. Nahant Bank (1840) at 166–7, that "[a] board of directors of the banks of Massachusetts is a body recognized by law. By the by-laws of these corporations, and by a usage, so general and uniform as to be regarded as part of the law of the land, they have the general superintendence and active management of all the concerns of the bank, and constitute, to all purposes of dealing with others, the corporation." It is hard to imagine a clearer rejection of the aggregate view. Similarly, in Hoyt v. Thompson's Executor, decided by the New York Court of Appeals in 1859, the court held that in corporate bodies the powers of the board of directors are, in a very important sense, original and undelegated. The stockholders do not confer, nor can they revoke those powers. They are derivative only in the sense of being received from the state in the act of incorporation. The directors convened as a board are the primary possessors of all the powers which the charter confers, and like private principals they may delegate to agents of their own appointment the performance of any acts which they themselves can perform. The recognition of this principle is absolutely necessary in the affairs of every corporation whose powers are vested in a board of directors. Without it the most ordinary business could not be carried on, and the corporate powers could not be executed. This constitutes a recognition that the aggregate view deriving from the membership corporation could not be maintained as a practical matter in corporations with hundreds or thousands of shareholders, as already existed in the 1850s.

17 See, e.g., the act adopted in 1837 by Connecticut permitting incorporation of "any lawful business," 1837 Conn. Pub. Acts 49, and various cases upholding such laws, e.g., Nesmith v. Sheldon (1849) at 817–19.

18 The same result was obtained in England by the adoption of the Regulation and Incorporation Act, 1856, ch. 63, § 2, 19 and 20 Vict. c. 47 (Eng.).

19 The transition from corporation to group of corporations occurred in the U.S. and Germany at the turn of the twentieth century (and continued during the fourth transformation hereinafter).

20 See also Mason v. Pewabic Mining Co. (1890) at 59 (stating that "we do not see that the right of the parties in regard to the assets of this corporation differ from those of a partnership on its dissolution").

21 This view was also reflected in contemporary books and law review articles. See, e.g., Freund 1897; Deiser 1908: 142; Laski 1916: 405–6; Machen 1911: 253 (all rejecting the aggregate view). Compare for a statement of the aggregate view. Morawetz 1882 ("the existence of a corporation as an entity, independent of its members, is a fiction").

22 Remarkably, this case involves a discriminatory state tax similar to the one struck down by Field on aggregate grounds in Santa Clara. See also similar statements in Ludwig v. W. Union Tel. Co. (1910) 163–4 (1910); Pullman Co. v. Kansas (1910) 64; W. Union Tel. Co. v. Kansas (1910) 36, which finally eliminated the restrictions imposed by Bank of Augusta v. Earl (1839). See Horwitz 1985: 177–87.

23 Another related development was the strengthening of limited liability resulting from the demise of the "trust fund" doctrine, which held that the capital stock of a corporation must be held in trust for the benefit of its creditors. This doctrine, which originated from Justice Story's opinion in Wood v. Dummer (1824), was upheld by the Supreme Court in Sawyer v. Hoag (1873). See also Cook 1887: 322. However, in 1892 the Supreme Court of Minnesota held in Hospes v. Northwestern Mfg. and Car Co., that "this trust fund doctrine … is not sufficiently precise or accurate to constitute a safe foundation upon which to build a system of legal rules … corporate property is not held in trust … Absolute control and power of disposition are inconsistent with the idea of trust. The capital of a corporation is its property … a corporation is in law as distinct a person as an individual is, and is entitled to hold

property (if not contrary to its charter) as absolutely as an individual can hold it." The doctrine then fell into desuetude, reinforced by the invention of no par stock in the early 20th century. See Horwitz 1985, *supra*.

24 Contemporary statement of the business judgment rule, see for example, ALI, Corporate Governance Project §4.01(c) (1994); Kamin v. American Express Co. (N.Y. 1976); Smith v. Van Gorkom (Del. 1985); Aronson v. Lewis (Del. 1984).

25 See also Manson v. Curtis (N.Y. 1918) at 323, in which the court held that "[d]irectors are the exclusive, executive representatives of the corporation and are charged with the administration of its internal affairs and the management and use of its assets. Clearly the law does not permit the stockholders to create a sterilized board of directors" (citations omitted).

26 See, e.g., Abbott v. Am. Hard Rubber Co. (N.Y. Sup. Ct. 1861), at 591–2.

27 See, e.g., People v. N. River Sugar Ref. Co. (1890), 626 (1890); State v. Standard Oil Co. (Ohio 1892), 184–5.

28 See also Cook 1894: 971–73 ("The courts are becoming more liberal, and many acts which fifty years ago would have been held to be ultra vires would now be held to be intra vires.") (emphasis omitted). By 1898 Cook wrote that "the doctrine of ultra vires is disappearing" Cook 1898.

29 On the "race to the bottom/race to the top" debate, see generally Bebchuk 1992; Cary 1974; Daines 2001; Winter 1977.

30 See also New Jersey Legislating for the United States, Indianapolis J., Nov. 11, 1901.

31 General Corporation Act of New Jersey sec. 51 (1890 rev.). See also sec. 104 (authorizing mergers).

32 Another significant development in this period was states passing statutes that allowed a majority of shareholders to sell corporate assets (before the 1890s, shareholder unanimity was required). This greatly facilitated mergers and also represented the decline of the aggregate view (Horwitz 1985: 201–2).

33 This state of affairs prompted Adolph Berle, the prime intellect behind the shareholder primacy doctrine in the 1930s, to concede defeat in 1954. See Berle (1954: 169): "Twenty years ago, the writer had a controversy with the late Professor E. Merrick Dodd, of Harvard Law School, the writer holding that corporate powers were powers in trust for shareholders while Professor Dodd argued that these powers were held in trust for the entire community. The argument has been settled (at least for the time being) squarely in favor of Professor Dodd's contention."

34 Fischel stated that "[a] corporation . . . is nothing more than a legal fiction that serves as a nexus for a mass of contracts which various individuals have voluntarily entered into for their mutual benefit" (Fischel 1982: 273). The point that the nexus of contracts theory is a reinvention of the aggregate view has been made repeatedly. See, e.g., Bratton 1989: 1471; Millon 1990: 229.

35 The real entity view is clearly the dominant one in sociology and some branches of economics. A whole branch of economic sociology centers on the study of organizations, and there are numerous books devoted to the topic. Most of these books revolve around the study of large corporations, since these are the dominant forms of organization in this society. See generally Powell and DiMaggio 1991; Pfeffer and Salancik 2003; Scott 2002; Thompson 2003; Smelser and Swedberg 1994, especially Part II, Section C, The Sociology of Firms, Organizations, and Industry. Moreover, they are informed by the economic perspective inaugurated by Coase in his classic "Nature of the Firm" article from 1937, and developed by Williamson and others into transaction cost economics. Coase 1937; Williamson 1994; for a critique see Granovetter 1985.

36 N.Y. Trust Co. v. Eisner (1921), 349.

37 "[O]ne person in law: as one person, they have one will, which is collected from the sense of the majority of the individuals . . . a person that never dies" Blackstone 1765: 456.

38 See note 4.

39 The corporate social responsibility literature portrays an optimistic reality in which corporations are in fact subject to some degree to corporate social responsibility monitoring, accounting and disclosure. See, e.g., Williams 1999: 1267–8, and Branson 2001: 641–4 (showing that large corporations are actually engaged in some form of social accounting, and referring to several social monitoring agencies, such as the Coalition for Environmentally Responsible Economics); Branson 2002a: 1219–22.

## Bibliography

### *Cases*

Abbott v. Am. Hard Rubber Co., 33 Barb. 578 (N.Y. Sup. Ct. 1861).
Aronson v. Lewis, 473 A.2d 805 (Del. 1984).
Bank of Augusta v. Earle, 38 U.S. (13 Pet.) 519 (1839).
Bank of the U.S. v. Dandridge, 25 U.S. (12 Wheat.) 64 (1827).
Beveridge v. N.Y. Elevated R.R., 112 N.Y. 1 (N.Y. 1889).
Burrill v. Nahant Bank, 43 Mass. 163 (1840).
Case of Sutton's Hospital, 77 Eng. Rep. 937 (1612) (KB).
De La Vergne Refrigerating Mach. Co. v. German Sav. Inst., 175 U.S. 40, 58 (1899).
Dodge v. Ford Motor Co., 170 N.W. 668 (Mich. 1919).
Dynamics Corp. of Am. v. CTS Corp., 794 F.2d 250 (7th Cir. 1986).
Hale v. Henkel, 201 U.S. 43 (1906).
Hospes v. Northwestern Mfg. & Car Co., 48 Minn. 174.
Hoyt v. Thompson's Executor, 9 N.Y. 207 (N.Y. 1859).
Kamin v. American Express Co. 54 A.2d 654 (N.Y. 1976).
Leslie v. Lorillard, 110 N.Y. 519 (N.Y. 1888).
Ludwig v. W. Union Tel. Co., 216 U.S. 146 (1910).
Manson v. Curtis, 223 N.Y. 313 (N.Y. 1918).
Marshall v. Baltimore & Ohio R.R., 57 U.S. (16 How.) 314 (1853).
Mason v. Pewabic Mining Co., 133 U.S. 50 (1890).
N.Y. Trust Co. v. Eisner, 256 U.S. 345 (1921).
Nesmith v. Sheldon, 48 U.S. (7 How.) 812 (1849).
Oregon Ry. & Navigation Co. v. Oregonian Ry., 130 U.S. 1 (1889).
Pembina Consolidated Co. v. Pennsylvania, 125 U.S. 181 (1888).
People v. N. River Sugar Ref. Co., 121 N.Y. 582 (1890).
Pullman Co. v. Kansas, 216 U.S. 56 (1910).
Revlon v. MacAndrews & Forbes Holdings, Inc., 506 A.2d 173 (Del. 1986).
Santa Clara Cty. v. S. Pac. R.R., 18 F. 385 (C.C.D. Cal. 1883).
Santa Clara Cty. v. S. Pac. R.R., 118 U.S. 394 (1886).
Sawyer v. Hoag, 84 U.S. 610 (1873).
Siegman v. Elec. Vehicle Co., 140 F. 117 (C.C.D. N.J. 1905).
Smith v. Van Gorkom, 488 A.2d 858 (Del. 1985).
South Ry. Co. v. Greene, 216 U.S. 400 (1910).
State v. Standard Oil Co., 49 Ohio St. 137 (Ohio 1892).
Tipling v. Pexall, 3 *Bulstrode* 233 (1614).
Trustees of Dartmouth College v. Woodward 17 U.S. (4 Wheat.) 518 (1819).
Union Pac. Ry. Co. v. Chicago, Rock Island & Pac. R.R., 163 U.S. 564 (1896).
United States v. E.C. Knight Co., 156 U.S. 1 (1895).
Unocal Corp. v. Mesa Petroleum Co., 493 A.2d 946 (Del. 1985).

West Union Tel. Co. v. Kansas, 216 U.S. 1 (1910).
Wood v. Dummer, 30 F. Cas. 435 (C.C.D. Me. 1824) (No. 17,944).

## *Statutes*

28 U.S.C. § 1332(c)(1) (2001).
1837 Conn. Pub. Acts 49.
C.Com. art. 33 (1807) (France).
General Corporation Law of New Jersey sec. 2 (1896).
General Corporation Act of New Jersey sec. 51 (1890 rev.).
New Jersey Legislating for the United States, Indianapolis J., Nov. 11, 1901.
Regulation and Incorporation Act, 1856, ch. 63, § 2, 19 & 20 Vict. c. 47 (Eng.).

## *References*


Allen, W.T. (1992), "Our Schizophrenic Conception of the Business Corporation," *Cardozo Law Review*, 14(2): 261–81.

American Law Institute, (1994), Principles of Corporate Governance, Analysis and Recommendations §2.01(b), §4.01(c).

Angell, J.K. and Ames, S. (1832), *A Treatise on the Law of Private Corporations Aggregate*, Boston: Hilliard, Gray, Little & Wilkins.

Arrow, K.J. (1973), "Social Responsibility and Economic Efficiency," *Public Policy*, 21: 303–17.

Avi-Yonah, R.S. (2002), "For Haven's Sake: Reflections on Inversion Transactions," *Tax Notes*, 95(12): 1793–9.

—— (2003), "National Regulation of Multinational Enterprises: An Essay on Comity, Extraterritoriality, and Harmonization," *Columbia Journal of Transnational Law*, 42(1): 5–34.

—— (2004), "Corporations, Society and the State: A Defense of the Corporate Tax," *Virginia Law Review*, 90(5): 1193–1255.

—— (2005), "The Cyclical Transformations of the Corporate Form: A Historical Perspective on Corporate Social Responsibility," *Delaware Journal of Corporate Law*, 30(3): 767–818.

Bebchuk, L.A. (1992), "Federalism and the Corporation: The Desirable Limits on State Competition in Corporate Law," *Harvard Law Review*, 105(7): 1435–1510.

Bebchuk, L.A. and Fried, J. (2004), *Pay Without Performance: The Unfulfilled Promise of Executive Compensation*, Cambridge: Harvard University Press.

Berle, A.A. (1931), "Corporate Powers as Powers in Trust," *Harvard Law Review*, 44(7): 1049–74.

—— (1932), "For Whom Are Corporate Managers trustees? A Note," *Harvard Law Review*, 45(8): 1365–72.

—— (1947), "The Theory of Enterprise Entity," *Colombia Law Review*, 47 (3): 343–58, reprinted in this volume.

—— (1954), *The 20th Century Capitalist Revolution*, New York: Harcourt, Brace.

Black, B. and Kraakman, R. (1996) "A Self-Enforcing Model of Corporate Law," *Harvard Law Review*, 109(8): 1911–82.

Blackstone, W. (1765), *Commentaries*.

Blair, M.M. (2003), "Locking in Capital: What Corporate Law Achieved for Business Organizers in the Nineteenth Century," *UCLA Law Review*, 51(2): 387–455.

Blair, M.M. and Stout, L.A, (1999), "A Team Production Theory of Corporate Law," *Virginia Law Review*, 85(2): 247–328.

Blumberg, P.I. (1986), "Limited Liability and Corporate Groups," *Journal of Corporation Law*, 11: 573–631.

—— (1993), *The Multinational Challenge to Corporation Law: The Search for aNew Corporate Personality*, New York: Oxford University Press.

Bradley, M., Schipani, C.A., Sundaram, A.K. and Walsh, J.P. (1999), 'The Purposes and Accountability of the Corporation in Contemporary Society: Corporate Governance at a Crossroads', *Law & Contemporary Problems*, 62(3): 9–86.

Branson, D.M. (2001), "Corporate Governance 'Reform' and the New Corporate Social Responsibility," *University of Pittsburgh Law Review*, 62(4): 605–47.

—— (2002a), "What is the New Corporate Social Responsibility? Corporate Social Responsibility Redux," *Tulane Law Review*, 76(5–6): 1207–26.

—— (2002b), "The Social Responsibility of Large Multinational Corporations," *Transnational Lawyer*, 16(1): 121–38.

Bratton, W.W. (1989), "The New Economic Theory of the Firm: Critical Perspectives from History," *Stanford Law Review*, 41(6): 1471–1527.

Business Roundtable (1990), *Corporate Governance and American Competitiveness*, Washington D.C.: Business Roundtable.

Cary, W.L. (1974), "Federalism and Corporate Law: Reflections on Delaware," *Yale Law Journal*, 83(4): 663–705.

Chen, R. and Hanson, J. (2004), "The Illusion of Law: The Legitimating Schemas of Modern Policy and Corporate Law," *Michigan Law Review* 103(1): 1–149.

Choi, S.J. and Guzman, A.T. (1998), "Portable Reciprocity: Rethinking the International Reach of Securities Regulation," *Southern California Law Review*, 71(5): 903–51.

Clark, R.C. (1986), *Corporate Law*, Boston: Little, Brown & Co.

Coase, R. (1937), "The Nature of the Firm," *Economica*, 4 (16): 386–405.

Coffee, J.C. (1990), "Unstable Coalitions: Corporate Governance as a Multi-Player Game," *Georgetown Law Journal*, 78: 1495–1549.

Cook, W.W. (1887), *A Treatise on the Law of Stock and Stockholders*, New York: Baker, Voorhis.

—— (1893), *The Corporation Problem*. New York: G.P. Putnam's Sons.

—— (1894), *A Treatise on Stock and Stockholders, Bonds, Mortgages and General Corporation Law*, 3rd edn, Chicago: Callaghan.

—— (1898), *A Treatise on the Law of Corporations Having a Capital Stock*, 4th edn, Chicago: Callaghan.

Daines, R. (2001), "Does Delaware Law Improve Firm Value?" *Journal of Financial Economics*, 62(3): 525–58.

Deiser, G.F. (1908), "The Juristic Person," *University of Pennsylvania Law Review*, 57(1): 131–42.

Dewey, J. (1926), "The Historic Background of Corporate Legal Personality," *Yale Law Journal*, 35(6): 655–73.

Dodd, E.M. (1932), "For Whom Are Corporate Managers Trustees?" *Harvard Law Review* 45(7): 1145–63.

Duff, P.W. (1938), *Personality in Roman Private Law*, Cambridge: Cambridge University Press.

Dwight, T.W. (1888), "The Legality of Trusts," *Political Science Quarterly*, 3(4): 592–632.

Easterbrook, F.H. and Fischel, D.R. (1981), "The Proper Role of a Target's Management in Responding to a Tender Offer," *Harvard Law Review*, 94(6): 1161–1204.

—— (1991), *The Economic Structure of Corporate Law*, Cambridge: Harvard University Press.

Fischel, D.R. (1978), "Efficient Capital Market Theory, the Market for Corporate Control, and the Regulation of Cash Tender Offers," *Texas Law Review*, 57(1): 1–46.
—— (1982), "The Corporate Governance Movement," *Vanderbilt Law Review*, 35(): 1259–92.
Fox, M.B. (1998), "The Political Economy of Statutory Reach: US Disclosure Rules in a Globalizing Market for Securities," *Michigan Law Review*, 97(3): 696–822.
—— (1999), "Retaining Mandatory Securities Disclosure: Why Issuer Choice is Not Investor Empowerment," *Virginia Law Review*, 85(7): 1335–1419.
Freund, E. (1897), *The Legal Nature of Corporations*, Chicago: Chicago University Press.
Friedman, M. (1970), "The Social Responsibility of Business is to Increase Its Profits," *New York Times Magazine*, Sept. 13.
Gierke, O. (1977), *Associations and Law: The Classical and Early Christian Stages*, George Heiman (trans.), Toronto: University of Toronto Press (originally published in 1881).
Graham, E.M. and Krugman, P.R. (1995), *Foreign Direct Investment in the United States*, Washington D.C.: Institute for International Economics.
Granovetter, M. (1985), "Economic Action and Social Structure: The Problem of Embeddedness," *American Journal of Sociology*, 91(3): 481–510.
Greenwood, D.J.H. (1996), "Fictional Shareholders: For Whom Are Corporate Managers Trustees, Revisited," *Southern California Law Review*, 69(5): 1021–1104.
Hansen, C. (1991), "Other Constituency Statutes: A Search for Perspective," *Business Lawyer* 46(4): 1355–76.
Hansmann, H. and Kraakman, R. (2000), "The Essential Role of Organizational Law," *Yale Law Journal*, 110(3): 387–440.
—— (2001), "The End of History for Corporate Law," *Georgetown Law Journal*, 89(2): 439–68.
Harris, R. (2000), *Industrializing English Law: Entrepreneurship and Business Organization, 1720–1844*, Cambridge: Cambridge University Press.
Harvard Law Review (2004), Developments in the Law – Corporations and Society, 117(7): 2169–295.
Hayek, F.A. (1960), "The Corporation in a Democratic Society: In Whose Interest Ought it and Will it Be Run," in M. Anshen and G.L. Bach (eds), *Management and Corporations*, New York: McGraw-Hill.
Holmes, O.W. (1897), "The Path of the Law," *Harvard Law Review*, 10(8): 457–78.
Horwitz, M.J. (1985), "Santa Clara Revisited: The Development of Corporate Theory," *West Virginia Law Review*, 88(1): 173–224.
Jensen, M.C. (2002), "Value Maximization, Stakeholder Theory, and the Corporate Objective Function," *Business Ethics Quarterly*, 12(2): 235–56.
Jensen, M.C and Meckling, W. (1976), "The Theory of the Firm: Managerial Behavior, Agency Costs, and Ownership Structure," *Journal of Financial Economics*, 3(2): 305–60.
Keasbey, E.Q. (1899), "New Jersey and the Great Corporations," *Harvard Law Review*, 13(3): 198–212.
King, A.A. and Lenox, M.J. (2000), "Industry Self-Regulation without Sanctions: The Chemical Industry's Responsible Care Program," *Academy of Management Journal*, 43(4): 698–716.
Larcom, R.C. (1937), *The Delaware Corporation*, Baltimore: Johns Hopkins Press.
Laski, H.J. (1916), "The Personality of Associations," *Harvard Law Review*, 29(4): 404–26.
Levitt, T. (1958), "The Dangers of Social Responsibility," *Harvard Business Review*, 36(5): 41–50.
Licht, A.N. (1998), "Regulatory Arbitrage for Real: International Securities Regulation in a World of Interacting Securities Markets," *Virginia Journal of International Law*, 38(3): 563–638.

Machen, A.W. (1911), "Corporate Personality," *Harvard Law Review*, 24(4): 253–67.

Manne, H.G. (1962), "The Higher Criticism of the Modern Corporation," *Columbia Law Review*, 62(3): 399–432.

Manne, H.G. and Wallich, H.C. (1972), *The Modern Corporation and Social Responsibility*, Washington D.C.: American Enterprise Institute for Public Policy Research.

Margolis, J.D. and Walsh, J.P. (2003), "Misery Loves Companies: Rethinking Social Initiatives by Business," *Administrative Science Quarterly*, 48(2): 268–305.

Mitchell, L.E. (1992), "A Theoretical and Practical Framework for Enforcing Corporate Constituency Statutes," *Texas Law Review*, 70(3): 579–644.

—— (ed.) (1995), *Progressive Corporate Law*, Boulder: Westview Press.

Millon, D. (1990), "Theories of the Corporation," *Duke Law Journal*, 1990(2): 201–62.

Morawetz, V. (1882), *A Treatise on the Law of Private Corporations Other than Charitable.*

Muchlinski, P.T. (1996), *Multinational Enterprises and the Law*, Boston: Little, Brown.

O'Connor, M.A. (1993), 'The Human Capital Era: Reconceptualizing Corporate Law to Facilitate Labor-Management Cooperation', *Cornell Law Review*, 78(5): 899–965.

Ostas, D.T. (2001), "Deconstructing Corporate Social Responsibility: Insights from Legal and Economic Theory," *American Business Law Journal*, 38(2): 261–99.

Pfeffer, J. and Salancik, G.R. (2003), *The External Control of Organizations: AResource Dependence Perspective*, Stanford: Stanford University Press (originally published in 1978).

Philips, M.J. (1994), "Reappraising the Real Entity Theory of the Corporation," *Florida State University Law Review*, 21(4): 1061–122.

Powell, W.W. and DiMaggio, P.J. (eds) (1991), *The New Institutionalism in Organizational Analysis*, Chicago: University of Chicago Press.

Orts, E.W. (1992), "Beyond shareholders: interpreting corporate constituency statutes," *George Washington Law Review*, 61(1): 14–135.

Reich, R.B. (1990), "Who is Us?" *Harvard Business Review*, 68(1): 53–64.

Reichert, A.K., Webb, M.S. and Thomas, E.G. (2000), "Corporate Support of Ethical and Environment Policies: A Financial Management Perspective," *Journal of Business Ethics*, 25(1): 53–64.

Romano, R. (1987), "The Political Economy of Takeover Statutes," *Virginia Law Review*, 73(1): 111–99.

—— (1998), "Empowering Investors: A Market Approach to Securities Regulation," *Yale Law Journal*, 107(8): 2359–430.

Rostow, E.V. (1961), "To Whom and for What Ends Is Corporate Management Responsible?" in E.S. Mason (ed.), *The Corporation in Modern Society*, Cambridge: Harvard University Press.

Schultz, F. (1951), *Classical Roman Law*, Oxford: Clarendon Press.

Scott, W.D. (2002), *Organizations: Rational, Natural, and Open Systems*, 5th edn, New York: Prentice-Hall (originally published in 1981).

Smelser, N.J. and Swedberg, R. (eds) (1994), *The Handbook of Economic Sociology*, Princeton: Princeton University Press.

Steffens, L. (1906), *The Struggle for Self-Government*, New York: McLure, Phillips & Co.

Thompson, J.D. (2003), *Organizations in Action: Social Science Bases of Administrative Theory.* New Brunswick, NJ: Transaction Publishers (originally published in 1967).

Tyson, L.D. (1991), "They Are Not Us: Why American Ownership Still Matters," *American Prospect*, 2(4): 37–49.

United Nations (2001), "Secretary-General Urges United States Business Leaders to Take Concerted Action Against "Unparalleled Nightmare" of AIDS," press release, U.N.

Doc. SG/SM/7827, AIDS/13 (June 1), available online at http://www.un.org/News/Press/docs/2001/sgsm7827.doc.htm/ (accessed 12 September 2006).

Vernon, R. (1998), *In the Hurricane's Eye: The Troubled Prospects of Multinational Enterprises*, Cambridge: Harvard University Press.

Walsh, J.P. and Avi-Yonah, R.S. (2005), "The Unfettered Corporation: Corporate Social Responsibility and the Coming Crisis of Corporate Control," unpublished manuscript (on file with author).

Wells, H.C.A. (2002), "The Cycles of Corporate Social Responsibility: An Historical Retrospective for the Twenty-First Century," *Kansas Law Review*, 51(1): 77–140.

White, J.B. (1999), *From Expectation to Experience: Essays on Law and Legal Education*, Ann Arbor: University of Michigan Press.

Williams, C.A. (1999), 'The Securities and Exchange Commission and Corporate Social Transparency', *Harvard Law Review*, 112(6): 1197–311.

—— (2002), "Corporate Social Responsibility in an Era of Economic Globalization," *University of California at Davis Law Review*, 35: 705–77.

Williamson, O.E. (1994), "Transaction Cost Economics and Organization Theory," in N.J. Smelser and R. Swedberg (eds), *The Handbook of Economic Sociology*, Princeton: Princeton University Press.

Winter, R.K. (1977), "State Law, Shareholder Protection, and the Theory of the Corporation," *Journal of Legal Studies*, 6(2): 251–92.

Wormser, I.M. (1912), "Piercing the Veil of Corporate Entity," *Columbia Law Review*, 12(6): 496–518.

# 11 The theory of enterprise entity*

*Adolf A. Berle Jr*

Classically, a corporation was conceived as an artificial person, coming into existence through creation by a sovereign power.

Thence proceeded certain advantages, which led the corporate form to become the principal method of organization of commercial, and especially of industrial, activity. Its primary business advantage, of course, was insulation of individual stockholders composing the corporation from liability for the debts of the corporate enterprise.[1] Realistic appraisal would today probably include a second advantage of weight: the distribution of responsibility for the enterprise among managing officials.[2] Many men are prepared to accept responsibility for a particular job in an enterprise who would not accept general responsibility for its entire functioning. Practically, in large-scale enterprises, few individuals can accept full management responsibility for an enterprise any more than they can accept full financial liability for its obligations.

As the scale of business enterprises enlarged, the process of sub-division began; hence subsidiary corporations wholly-owned or partly-owned; or holding companies combined into a series of corporations constituting a combined economic enterprise; and so forth. More often than not, a single large-scale business is conducted, not by a single corporation, but by a constellation of corporations controlled by a central holding company, the various sectors being separately incorporated, either because they were once independent and have been acquired, or because the central concern, entering new fields, created new corporations to develop them, or for tax reasons. In some instances, departments of the business are separately incorporated and operated as separate legal units.[3] Since under modern corporation statutes any three persons can in effect request and get a corporate charter, writing their own grant of powers (with very few limitations) and thus constituting themselves an artificial person, the process has been easy to carry on, and has been a decided business convenience.

This is far from the original conception of a corporation. The legal doctrine of corporate personality was built around the idea of a sovereign grant of certain attributes of personality to a definable group, engaged in an enterprise. The so-called "artificial personality" was designed to be the enterpriser of a project. Multiplicity of artificial personalities within an enterprise unit would probably have been impossible under most early corporation laws.[4]

The divergence between corporate theory and the underlying economic facts has occasioned a variety of problems (dealt with *ad hoc* by the courts) in which the theory of "artificial personality" simply did not work, and was consequently extended, disregarded, sometimes buttressed by further fiction, at others manipulated to get a convenient result.

All this suggests that a review of the classic conception is in order. It has seemed to the writer that one of the pressing needs in the field of corporation law is its systematization. A series of rules have been adopted in varying groups of cases as wrongs appeared and remedies were worked out. These emerged as isolated doctrines applicable to specific situations. This essay is designed to suggest the possibility that a number of rules which are regarded as separate in fact are applications of a single dominant principle.

The theory of enterprise entity is capable of application in many fields of corporation law besides those here examined.

It is the thesis of this essay:

- That the entity commonly known as "corporate entity" takes its being from the reality of the underlying enterprise, formed or in formation;
- That the state's approval of the corporate form sets up a *prima facie* case that the assets, liabilities and operations of the corporation are those of the enterprise;
- But that where the corporate entity is defective, or otherwise challenged, its existence, extent and consequences may be determined by the actual existence and extent and operations of the underlying enterprise, which by these very qualities acquires an entity of its own, recognized by law.
- For brevity, this hypothesis is hereafter referred to as the theory of "enterprise entity."
- It is believed that application of this theory can systematize the scattered rules of corporation law in a number of areas.
- Examination of this theory (which could be applied in other branches of corporation law) is here made in respect of three main fields: the field commonly called "*de facto*" corporations; the field commonly comprehended by "disregard of the corporate fiction" (which is a name applied to a multiplicity of diverse situations); and the field of added liability of shareholders opened by the recent case of *Anderson v. Abbott.*[5]

Review of the law of corporate entity and of the exceptions made to it, discloses a steady and growing number of instances in which, on one pretext or another, the courts have either

(a) erected corporate personality which the state had not granted; or
(b) disregarded corporate personality where the state had granted it,

for the real purpose (in both categories) of giving legal effect to factual relationships set up between an economic entity and an outsider. The various

reasons, fictions, arguments and important considerations are many, diverse, and frequently inconsistent; but the scheme of these various exceptions is none the less consistent and logical enough. The corporation is emerging as an enterprise bounded by economics, rather than as an artificial mystic personality bounded by forms of words in a charter, minute books, and books of account. The change seems to be for the better.

A variety of situations may be summarized in which the courts grant or deny the existence of corporate entity where the legal entity does not exist; or deny the existence of a corporate entity where the legal entity undoubtedly does exist. These perhaps outline the emerging doctrine. A like summary of situations, in which corporate action is sustained on the basis of the needs of an economic entity when it could not easily be sustained on the logic of cold legal entity, tends to fill in the picture.

## Erection of entity where legal entity does not exist

### *"De Facto" corporation*

The simplest and most direct disregard of the theory of state-created artificial entity occurred when courts developed the doctrine of "*de facto*" corporations. Here, quite plainly, a group of people associated in an enterprise have not secured creation of the entity because they have failed to do some act or to fulfil some condition essential to state action. Despite this fact, courts struggled in a variety of ways to attach to the acts of this imperfectly clothed enterprise the consequences which would have attached had it been validly incorporated, and the doctrine of "*de facto*" corporations emerged. The main sequences are these:

- The enterprise contracts with an outsider, who later brings action against the enterprise as though it were a corporation; and the enterprise is held liable in corporate form.[6]
- The enterprise contracts with an outsider, and subsequently brings action in corporate form against the outsider. The outsider is held liable to the enterprise.[7]
- The enterprise contracts with an outsider, and the outsider brings action against the component individuals. They are absolved from liability and the outsider held to his remedy against the enterprise only.[8]
- The enterprise contracts with an outsider, and the component individuals seek to hold the outsider liable on his contract. It would seem to follow logically that the individuals are not allowed to recover: recovery must be by the enterprise.

In the first of these four situations, courts frequently invoke the doctrine of estoppel *in pais* – that is, that the enterprise has represented itself to be a corporation, that the outsider has acted on that assumption, and that the enterprise cannot now be heard to deny its corporate existence. This is, or at least can be,

true estoppel; and, by consequence, these cases do not necessarily involve the creation, by court law, of an artificial entity – though they undeniably have that effect.

But the second and third situations do involve court-law setting up an artificial entity despite the fact that the state has not done so. The decisions not infrequently say that the outsider is "estopped" to deny that the *de facto* corporation plaintiff is not an actual corporation. It is fairly plain that this is the use of a word rather than a rule; the court, for reasons it deems sufficient, is bluntly telling the outsider that he will be held to his obligation with this particular group as an entity despite the fact that its entity was not created as the state prescribed. In the third situation even the language of estoppel is often discarded, and the general rule is stated as being that where the component individuals have endeavored to form a corporation, have, in good faith, believed that they were such a legal entity, and have colorably complied with the incorporation law, they will be protected from individual liability; and, equally, they have lost their power to enforce liabilities against outsiders in favor of themselves as individuals.

Various arguments are used to reach this result, and most of them are highly persuasive of the proposition that the dealings of this enterprise, though defectively incorporated, should produce the same results in the respective cases as though it were well, fully and truly incorporated. There is the argument of business convenience; and the argument that it would be inequitable to impose individual liability on associates where they did not intend to create it and the outsider did not seek to have it; and so forth. But these arguments can hardly be said to establish the proposition that, because of these facts, a set of arrangements which the statutes say did not create a separate personality had somehow created one, or at least produced some of its consequences.

What they do suggest is something quite different. They suggest that an agreement of the associates, made in such a way as to cause them reasonably to believe that they are a corporation, accompanied by enterprise action (user) has created an economic fact, namely, a composite unit. Probably the most essential element, in this view, is that of "user" – the element of an actual economic fact.[9] If we were disregarding the historical language, and the fictions by which courts habitually score an advance from one point to another, we should say something like this: Here a group of individuals have agreed with each other that they will carry on an enterprise as a unit, limiting their liability to stated contributions of capital. So far as entity could be created by agreement, and liability limited by agreement, they did it. The group thus created embarked on the intended enterprise. Outsiders accepted this enterprise-fact and dealt with it. So far as economics and agreement between the parties was concerned, the unit was complete. Unhappily, for their failure to conform to some condition, the state did not sanctify the union. In effect we have something like a common-law marriage. The union is not contrary to any policy of the state; the "colorable" compliance, and the fact that (if the associates had been more careful) they

could have had the state's blessing sufficiently indicate that. In consequence, the acts of the entity have the same result and effects as would corporate acts – subject, of course, to possible attack by the state, but not otherwise. Though the words "*de facto* corporation" are somewhat slighted by modern commentators,[10] they state the real case.

### *Combinations of corporations into an aggregate enterprise entity*

Another illustration of judicial erection of a new entity occurs in situations where the corporate personality (as embodied in its charter, books and so forth) does not correspond to the actual enterprise, but merely to a fragment of it. The result is to construct a new aggregate of assets and liabilities. Typical cases appear where a partnership or a central corporation owns the controlling interest in one or more other corporations, but has so handled them that they have ceased to represent a separate enterprise and have become, as a business matter, more or less indistinguishable parts of a larger enterprise. The decisions disregard the paper corporate personalities and base liability on the assets of the enterprise. The reasoning by which courts reach this result varies: it is sometimes said that one corporation has become a mere "agency" of another; or that its operations have been so intermingled that it has lost its identity; or that the business arrangements indicate that it has become a "mere instrumentality."

The courts frequently decide that, by so dealing with a subsidiary, the parent corporation becomes liable for the obligations of the subsidiary, which, of course, has not ceased to be liable itself. In most such cases the discussion goes chiefly on the negative point of disregarding the corporate fiction, which would assign separate existence to the subsidiary, and insulate the stockholder from its liabilities and obligations. Of more importance here is the positive corollary: liabilities attached to the assets and operations of the parent stockholder as well as to those of the fragment. In general, in this type of case (parent and subsidiary) liability comes to rest there; and the result is that the assets of parent and subsidiary taken together are affected with the obligations of one of the component legal personalities.[11]

This category of cases stands still more squarely on the foundation of economic enterprise-fact. The courts disregard the corporate fiction specifically because it has parted company with the enterprise-fact, for whose furtherance the corporation was created; and, having got that far, they then take the further step of ascertaining what is the actual enterprise-fact and attach the consequences of the acts of the component individuals or corporations to that enterprise entity, to the extent that the economic outlines of the situation warrant or require.

Quite early in the proceedings a partnership, creating a corporation, and using it as an instrumentality (with appropriate circumstances showing that the partners themselves disregarded the corporate enterprise) was disregarded and the consequences of the corporate acts attached to the partnership.[12] Where

several corporations became in effect a single enterprise and merged their operations, their several entities were disregarded and their respective assets and liabilities were pooled in a common pot which represented, substantially, the actual enterprise of which they were parts.[13] While it is clear that mere ownership of all of the stock of one corporation by another does not merge the enterprises if, in point of fact, they are separately maintained and separately operated,[14] yet something more than words in corporate minute books and separate figures in account books are needed. In one line of cases, the fact that the sole owner of stock in a corporation has embarked the concern in large operations with no adequate capital or assets to support such operations, is held to justify the conclusion that this insufficiently outfitted entity did not really conduct a separate enterprise, and the parent corporation is held liable for its obligations.[15] Some of these cases are particularly interesting from the technical standpoint: for they suggest that to preserve the independence of an enterprise which is needed to support the continuance of separate legal personality, the stockholders must provide the entity with separate assets sufficient to give it at least a reasonable business chance to carry out its asserted functions. In these, as in the cases of *de facto* corporations, the courts' rulings construct a new entity, this time out of spare parts distributed among component corporations. But they go further; and after disregarding the fictitious personality where it does not correspond with the enterprise, they outline an entity with a body of assets to which liabilities are assigned more nearly in accord with the ascertainable fact of the enterprise and its relationship to outsiders.[16]

In effect what happens is that the court, for sufficient reason, has determined that though there are two or more personalities, there is but one enterprise; and that this enterprise has been so handled that it should respond, as a whole, for the debts of certain component elements of it. The court thus has constructed for purposes of imposing liability an entity unknown to any secretary of state comprising assets and liabilities of two or more legal personalities; endowed that entity with the assets of both, and charged it with the liabilities of one or both. The facts which induce courts to do this are precisely the facts which most persuasively demonstrate that, though nominally there were supposed to be two or more enterprises, in fact, there was but one. The economic fact pushes through the paper differentiations embodied in the corporate certificates; and liabilities are dealt with in accord with the business, instead of the legal fact of corporate entity.

It is no long step from this to the doctrine of the Deep Rock case,[17] and sundry decisions like it. There, the decision did not merely recognize a combination of separate legal personalities in the operation of a single enterprise, but undertook to reconstruct the financial position of various participants. Thus, a debt due the parent from the subsidiary was postponed and made junior to the rights of outside stockholders in that subsidiary: the idea being, in substance, that in the combined enterprise the contribution made by the parent, although in form a loan to the subsidiary, was in fact an equity investment. Judicial recognition of an entity according to economic fact is thus developed by

determining the position of the various claimants in accordance with this fact: the judicial creation of a scheme of finance.

We may conclude that in an apparently growing number of situations[18] the courts, in effect, mold the corporate situation to the economic fact; that the economic fact is the actual business enterprise as carried on by the component individuals active in it; and that the entity, with its attendant consequences of a particular body of assets and operations, is then given legal attributes which would have been given to it had it been a body corporate duly and properly enfranchised by the state.

### *Special liability enterprises*

A third category of cases is just appearing, and it pushes the doctrine into new ground. These are cases in which the enterprise is itself legal and violative of no policy or rule of law; but additional liabilities, over and above the stockholders' contribution to capital are imposed by law.

In Anderson v. Abbott[19] a group of individuals wished to purchase the shares of certain banks and for this purpose caused the Banco Kentucky Corporation to be formed and to invest the bulk of its capital in a majority stock interest in seven banks and a minority interest in an eighth. There were other investments; but Banco was plainly formed for the purpose of controlling these banks, a "plain indication of the nature of the enterprise" being given (as the court later pointed out) when Banco so stated in its application to list its shares on the Chicago Stock Exchange. One of the banks it controlled failed; the bank's receiver obtained judgment against Banco for the double liability imposed on bank shareholders, but could not collect in full; and the receiver brought action against the individual stockholders of Banco to collect the balance of the assessment. If Banco was the bank's stockholder, and the statutory liability ran only against it, the receiver had no case. The circuit court of appeals so held; but was reversed by the Supreme Court, Mr. Justice Douglas remarking: "We are dealing here with a principle of liability which is concerned with realities, not forms. As we have said the net practical effect of the organization and management of Banco was the same as though the shares of the Bank were held in trust for beneficiaries who were in point of substance its only owners. Those who acquired shares of Banco did not enter upon an enterprise distinct from the banking business."[20] Accordingly, a majority of the Court held the individual stockholders of Banco liable for the balance of the unpaid double-liability assessment of the bank stock which Banco itself was unable to satisfy.

This was a stern analysis of Banco as an enterprise, the majority of the Court reaching the conclusion that it was in effect a trusteeship of shares affected with a double liability held in trust for its shareholders. Given this conception of the nature of the enterprise, the decision was logical. Possibly the same result could have been reached by a different fact analysis which might have freed Justice Douglas from a somewhat far-fetched reconstruction of the Banco enterprise as a kind of voting trust. The Court might have determined that the enterprise of

Banco was that of being a stockholder in banks (which was the fact); that this kind of enterprise necessarily involved, not only putting up money to buy bank shares, but also assumption of double liability on those shares should the bank fail; and that accordingly the stockholders in Banco undertook this as a liability of the enterprise in which they were engaged when they purchased their stock in Banco. It would then be logical, procedurally, to permit a receiver or trustee in bankruptcy of Banco to collect from Banco's stockholders so as to enable Banco to fulfill the obligations the enterprise had marked out for itself.[21] The dissent of four justices in the case was based on the proposition that the majority gave undue vigor to the legislative policy of double liability which did not justify disregard of the corporate fiction. The essential difference between Mr. Justice Douglas and Mr. Justice Jackson, who wrote the dissenting opinion, seems to be in their differing analysis and reconstruction of the enterprise – in that regard, a difference of opinion was possible – but, granted the economic analysis of the majority, the decision reached is defensible.

## Refusal to recognize corporate entity where the enterprise has become objectionable

The cycle is completed by reference to classic cases in which the fiction of corporate organization has been disregarded, but without recreating or recognizing any other entity. Here the consequences of corporate organization are not allowed to operate; and individuals or component elements are held responsible for criminal penalty, personal liability, or other legal results. This is a well recognized chapter of the law of corporations; and it is so well known that little more than index reference to it is needed here. Courts have long recognized that, despite its long history of entity, a corporation is at bottom but an association of individuals united for a common purpose and permitted by law to use a common name. The literature of disregard of corporate fiction where the corporation was merely in aid of or adjunct to an objectionable enterprise is voluminous. The fragile quality of the legal personality created by a corporation is aptly demonstrated. When the corporate fiction is disregarded, an actual underlying enterprise entity may be made to appear. The cases, negative in quality, which disregard the corporate fiction and hold individuals or component elements liable, seem to stem from the same principle – except that in such cases the courts hold that the enterprise itself, with or without corporate clothes, is objectionable and the individuals must be held criminally or civilly accountable. A summary demonstrates the point.

Notable among the situations in which the corporate fiction has been disregarded are those in which the individuals or elements composing a validly created corporation are really endeavoring to violate some criminal law,[22] perpetrate some fraud,[23] or contravene some settled and defined rule of policy.[24] Such are cases in which the challenged corporation has been held to be a screen behind which the individual stockholders or associates design to violate the antitrust laws; or statutes forbidding foreign ownership of American flag vessels; or

laws forbidding alien enemy ownership of property. Violation of policy, if not of criminal law, produces the same effect: hence the corporate fiction can be disregarded where individuals endeavor to use a corporation to screen a transfer in fraud of creditors. Instances are legion. In all these cases, negative though they are, the point is that the enterprise, however conducted, is forbidden by law. The illegal element, of course, does not appear in the certificate of incorporation; but the facts show it to be present. In effect, the courts look through the paper delineation to the actual enterprise; and then determine whether it is criminal, illegal, contrary to public policy, or otherwise bad (as the circumstances may be) for individuals to conduct that enterprise by any kind of organization.

It is highly probable that the courts would have no great difficulty even if the illegality were directly authorized by the incorporation papers. The Ohio Corporation Law, for instance, empowers a corporation to do anything which "natural persons" could do. A certificate of incorporation might be filed which stated as one of the powers and purposes that the concern proposed to commit criminal acts. Yet if such a document should escape rejection by the Secretary of State, there is little doubt that it would not serve even as a *prima facie* defense to criminal action against the shareholders. In fact, the courts have repeatedly examined the underlying enterprise to find what its real purport was: for example, to act as a channel for illegal rebates,[25] or to combine railroads in violation of the anti-trust laws;[26] and indeed courts will inquire into this in almost any case in which violation of law was an object of the enterprise.

The foregoing analysis suggests that the variety of rules formulated in individual cases and designed to effect remedies in specific situations can now be grouped and systematized. We should no longer have a specific rule of subordination (as in the Deep Rock case) because a parent has dealt badly with a subsidiary; or a rule that a holding company owning bank shares does not isolate its stockholders from double liability on the bank stock; or a rule that a corporation may not act as a cloak for an operation designed to defraud creditors, and so on. Rather we have a principle of law, namely, that below the corporation papers there is always an enterprise; that, *prima facie*, the corporate papers, minute books and books of account describe and indicate that enterprise; but that it is always open to inquiry whether the enterprise-fact corresponds to the corporate-fact.[27] If it does not, the court may, on cause shown, insist upon dealing with the actuality of the enterprise-fact, and may impose the remedy which corresponds to the fact. Even in the very recent litigation of Zahn v. Trans-America Corporation,[28] it could have been argued without too much difficulty that the real vice of the corporation there attacked was the creation by directors representing one group of stockholders of a separate enterprise inside of and adverse to the remaining stockholders, instead of operating the enterprise for the benefit of all, as the corporation papers contemplated.

In all these categories, the underlying principle seems plain. Whenever "corporate entity" is challenged, the court looks at the enterprise. Where the enterprise as such would be illegal or against public policy for individuals to conduct, that enterprise is equally illegal when carried on by a corporation, and the corporate

form is not a protection. This is, in essence, not so much a "disregard of the corporate fiction" as it is a holding that the economic enterprise, whether corporate or non-corporate, is illegal, or criminal, or in violation of public policy, or fraudulent, or otherwise objectionable, as the case may be. The nature of the enterprise determines the result, negativing the corporate personality or any other form of organization of that enterprise.

If it be shown that the enterprise is not reflected and comprehended by the corporate papers, books and operation, the court may reconstruct the actual enterprise, giving entity to it, based on the economic facts. Thus one corporation may be shown to be in fact only an "instrumentality" of a larger enterprise, or to be so intermingled with the operations of such larger enterprise as to have lost its own identity. On such reconstruction of the true entity, the court may assign the liabilities of the paper fragment to the economic whole; or may (as in the Deep Rock case) assign priority or subordination to its liabilities or securities stock, so as to attain, as nearly as possible, a financial result corresponding to the reasonable expectations of the creditors and security holders.

In a new category of cases, forecast by the Anderson decision, where again the corporate entity is challenged, if the enterprise is such as to entail liability over and above the usual stockholders' contribution of capital, added liability may be imposed on the component elements constituting it.

## Some further implications

A substantial vista of speculative thinking is suggested by this analysis. Some illustrations may prove of interest.

Application of the theory of enterprise entity would contribute to solving a number of problems. For example, in Cintas v. American Car and Foundry Company[29] the plaintiff sought payment of a preferred dividend in priority to payment of a common dividend. Under the law of New Jersey, if the dividend he sought had been earned, he was entitled to this priority; and the problem was whether it had in fact been earned. The principal defense was that, while the corporation in which plaintiff held preferred stock had shown earnings on its income statements, these were in fact more than counter-balanced by losses suffered through the corporation's subsidiary; and that on consolidation the earnings forming the basis of plaintiff's claim would disappear. The court clung to the corporate entity, and the plaintiff got his decree, though the enterprise as a consolidated whole (on the facts shown) had demonstrably not earned the dividends in question. The doctrine of enterprise entity would have led to a healthier result; for, realistically, on a basis of economic fact, the plaintiff was not entitled to his dividends, and indeed it would probably have been illegal for the corporation to pay them.

Clearly, also, whenever an enterprise is composed of more than one corporate entity, two distinct sets of relationships are entailed. The first consists of a body of relationships which the enterprise has with individuals – outside stockholders and outside creditors of the various corporations involved. A second and wholly

different set of relationships exist by reason of the distribution of liabilities or security holdings within the enterprise – intra-subsidiary debts, stockholdings, contract liabilities, and so forth. There can be no doubt at all about the economic differentiation. Where, for instance, a holding company has a majority stock interest in half a dozen subsidiaries, the shares of the subsidiary corporation held by the parent stand in a different class from the shares held by outsiders. The outsider is participating in one enterprise, that of the subsidiary, and his interest lies solely with it. The parent or intra-subsidiary holdings, however, form a composite group aggregated into a larger and quite dissimilar enterprise. The directors of the parent, for example, considering the parent's shareholders, might be entirely wise and faithful in causing the expansion of Subsidiary X into a profitable field, while restraining Subsidiary Y from entering that territory. Presumably they do this because Subsidiary X is better fitted for the work, and the profits accruing to the whole or aggregate enterprise will be greater. But this would be little comfort to an outside stockholder in Subsidiary Y, who saw his fragment of the entire enterprise languishing, while another fragment became extremely profitable.[30]

Even more striking is the difference between outside creditors as against intra-enterprise or inter-corporate debts. When Subsidiary X owes money to an outside creditor, the transaction is at arm's length and the debt has to be paid. But if Subsidiary X borrows from or loans to its parent or a brother subsidiary, it will be able to collect, or be required to pay, solely depending on the business interests of the aggregate enterprise – a quite different set of economic circumstances and motivations.

The clearest picture given in these cases is that appearing in the consolidated balance sheet. Here, an accountant undertakes to reconstruct precisely the aggregate enterprise, and to set up the corresponding figures. Few will question that the results shown by a consolidated balance sheet often differ markedly from the picture given by the separated accounts of the component corporate parts of such an enterprise. The accountant is not imprisoned by the technical outlines of artificial intra-corporate personality. He can and does draw a financial picture of the entire enterprise; and in so doing he differentiates between the financial relationships of that enterprise with those outside it and the distribution of financial interest and liability as between its interior component parts. If the accountant's work is well done, the result is, or should be, a statement of financial facts applicable, not to any legal entity, but to an enterprise entity.

By that test, certain agreements made by a subsidiary which would normally be outside the pale of propriety may well be justified on the theory of enterprise entity. A holding company, for instance, issues its debentures. The buyers of these, being financially wise, know perfectly well that the debts of the subsidiary can be collected out of assets of the subsidiary, diminishing the assets contributable to the stock which is the property of the holding corporation and are, therefore, senior to the holding company's debentures. In consequence, the buyers of these debentures exact an agreement[31] from the holding company that it will not permit its subsidiary corporations to mortgage their assets or create an

indebtedness beyond a certain amount, except to the parent company. This is a proper protection for the new debenture holders. From the point of view of an outside shareholder or creditor of the subsidiary this may well preclude the subsidiary from obtaining additional financing which it might badly need. But, if the enterprise as a whole were analysed, it might well prove that sacrifice by the parent of the subsidiary's freedom of action was more than counter-balanced by advantages to the subsidiary. For instance, a telephone company may have many subsidiaries who have derived advantage not only from their own local operations but from the thought that they have access to the vast system of mainline communications operated by the parent. This benefit may more than compensate the disadvantage which the subsidiary incurs by subordinating its finances to the general pattern of the parent; and upon analysis the entire enterprise might equally find itself affected with a duty to finance its subsidiary, or, in case of failure, to respond for its subsidiary's debts.

This is not at all to say that in a case in which one corporation owns a controlling stock interest in another it has by that fact alone made a single enterprise out of two different ones. The controlling corporation has a choice. It can, if it chooses, elect to permit, or perhaps require, its subsidiary to manage its own affairs, make its own decisions, and operate as a separate enterprise, the parent retaining only an investor's interest. Or it can integrate the operations of the subsidiary with its own, in whole or in part, thereby bringing the two operations together into a single enterprise entity.[32] There is no compulsion on it to adopt or refrain from either course; but the legal consequences vary with the choice.

Where in these cases the separate entity of the enterprises as well as their separate corporate personality has been preserved, there is no need of departing from the paper organization – if the paper organization corresponds to the fact. But where the interests of one of the corporate units in the composite picture have been sacrificed to the legitimate interests of the entire enterprise, it might be found by way of remedy that the outside stockholder in the sacrificed subsidiary can be most fairly treated by giving to him a position equal to the outside stockholder in the parent since, in the enterprise-fact, his contribution as stockholder has been used for the general benefit of the entire enterprise, and to the detriment of his position as a stockholder in a fragment of the enterprise. This indeed will not be any very great extension of the Deep Rock doctrine. There an intra-corporate debt was subordinated to stockholders of the subsidiary since it was in effect found to be a capital contribution. In the case proposed, the stockholder in a subsidiary might be found to have been an enterprise-fact, a contributor to capital in the entire enterprise and entitled to share as such.

The theory of enterprise entity could likewise systematize other unnecessarily rough edges in the present law of corporations. There is, for instance, the technically accurate but economically unsound rule that no corporation can be made to pay for services rendered in advance of its organization, except by its specific consent. Naturally, when the artificial personality does not exist it can incur no liability. Yet, economically, enterprises must be assembled before they can be organized as a corporation, and in fact quite usually the enterprise is the

result of the work of promoters, lawyers, and perhaps others, carried on for some period of time before the incorporation papers are finally approved. The work of these people enters into, and indeed is, the body of the enterprise when it is first formed. For the enterprise, and not the incorporation papers, is the true entity, and there should be no difficulty in charging to the corporation when formed the reasonable value of the services rendered with due regard being had to the fact that, in this early period, the organizers owe the highest duty and good faith in determining the amount of such liability since they are virtually dealing with themselves. It was absurd to say, as one court did,[33] that a lawyer who incorporated a company could collect a bill for holding the incorporators' meeting, but could not include in his bill the time spent working with the incorporators in developing the general corporate plan. It is equally unreal to assume that promoters work for nothing, and that their services are worthless because rendered before the date of incorporation. To reach this result by dealing frankly with the fact that an enterprise is itself an entity, and is recognized as such in business fact, avoids some inequitable results determined by the courts, and, still more important, makes unnecessary the more or less recognized subterfuges and evasions of technical rules so as to permit the enterprisers to get a start.

Other illustrations of possible systematization of corporate law along this line could be given. Those set out perhaps sufficiently indicate the need for increased study along this line.

## Notes

* Adapted from *Columbia Law Review*, 47 (3), April 1947, pp. 343–58.

1 For a discussion, see Douglas and Shanks, "Insulation from Liability through Subsidiary Corporations," 39 *Yale L. J.* 193 (1939).

2 For a discussion, see Drucker, Concept Of The Corporation (1946).

3 See, e.g., Registration Statement under Securities Act of 1933 (covering Indenture, dated as of April 1, 1947, securing the Thirty-Five Year Debentures of American Telephone and Telegraph Co. due April 1, 1982). Thus, Bell Telephone Laboratories, Inc. is owned 50% by American Telephone and Telegraph Co., and 50% by Western Electric Corporation, which in turn carries on the business of manufacturing apparatus primarily for American Telephone and Telegraph Company which owns the stock of Western.

4 In the writer's opinion, the best study is still Taylor, Corporations (5th edn 1902), cc. 1, 2. Though the book is long out of date, Taylor was an historian as well as a lawyer, and his historical review has not yet been seriously modified. See also Timberg, "Corporate Fictions: Logical, Social and International Implications", 46 *Columbia Law Rev.* 533 (1946). Many early corporation laws in this country confined corporations to a single enterprise, though the "single purpose rule," as it was called, proved unworkable as business organizations evolved.

5 321 U. S. 349 (1943).

6 See, e.g., Frawley v. Tenafly Trans. Co., 95 N. J. L. 405, 113 A.2d 242 (1920); Empire Mfg. Co. v. Stuart, 46 Mich. 482, 9 N. W. 527 (1881); Dooley v. Cheshire Glass Co., 15 Gray 494 (Mass.1860).

7 Commercial Bank v. Pfeiffer, 108 N. Y. 242, 215 N. E. 311 (1888); Ingle System Co. v. Norris and Hall, 132 Tenn. 472, 178 S. W. 1113 (1915); Societe Titanor v. Paxton and Vierling Iron Works, 124 Neb. 570, 247 N. W. 356 (1933).

8 Inter-Ocean Newspaper Co. v. Robertson, 296 I11. 92, 129 N. E. 523 (1921); Whitney v. Wyman, 101 U. S. 392 (1872); 1 Cook, Corporations § 234 (4th edn 1898).

9 Many of the cases bear heavily on use of the corporate name, or a name suggesting corporate existence. Lowell-Woodward Hardware Co. v. Woods, 104 Kan. 729, 180 Pac. 734 (1919); *cf.* Societe Titanor v. Paxton and Vierling Iron Works, 124 Neb. 570, 247 N. W. 356 (1933). But the thrust of the argument in substance is that the associates did in fact act like a corporation, the use of the name being persuasive evidence of that fact.

10 The term "defectively organized corporations" is used throughout one of the leading casebooks (Dodd and Baker1940), following a terminology suggested by Professor Burdick. Burdick, "Are Defectively Incorporated Associations Partnerships?" 6 Columbia Law Rev. 1 (1906).

11 For discussion of the various theories, see Latty, E. R., *Subsidiaries And Affiliated Corporations* (1936); Rembar, "Affiliated Companies Claims," 39 *Columbia Law Rev.* 907, 923 (1939); Douglas and Shanks, "Insulation from Liability through Subsidiary Corporations," 39 *Yale L. J.* 193 (1929); Hornstein, "Legal Controls for Intra-Corporate Abuse,"41 *Columbia Law Rev.* 784 (1941); Israels, "Implications and Limitations of the 'Deep Rock Doctrine'," 42 *Columbia Law Rev.* 376 (1942).

12 *In re* Rieger, 157 Fed. 609 (S. D. Ohio 1907), the first of a long line of cases.

13 Sampsell v. Imperial Paper Co., 313 U. S. 215 (1940). The corporation was found to be a mere "pocket" of its majority stockholders. Its creditors were forced to share in the assets *pari passu* with the stockholder's creditors.

14 *In re* Lawyers' Mortgage Co., 284 N. Y. 371, 31 N. E.2d 492 (1940); Albright *v.* Jefferson County Nat. Bank, 292 N. Y. 31, 53 N. E.2d 753 (1944) (citing cases).

15 In all cases insufficient capitalization is persuasive evidence that the enterprise was not separate. See Israels, *supra* (note 11), at 381, 382. Luckenbach S. S. Co. v. Grace & Co., 267 Fed. 676 (C. C. A. 4th 1920); Oriental Investment Co. v. Barclay, 25 Tex. Civ. App. 543, 64 S. W. 80 (1901); *cf.* Richards v. Mayfair, Inc., 287 Mass. 280, 191 N. E. 430 (1934); Erikson v. Minnesota & Ontario Power Co., 134 Minn. 209, 158 N. W. 979 (1916); Pepper v. Litton, 308 U. S. 295 (1940); Taylor v. Standard Gas & Elect. Co., 306 U. S. 307 (1939); *Inre* Pittsburgh Rys., 155 F.2d 477 (C. C. A. 3d 1946).

16 The current discussions are chiefly based on the fact that, in these cases, a claim of the parent against the subsidiary may be subordinated to claims of outside creditors, or even stockholders, of the subsidiary. Yet, clearly, admission of a subsidiary's creditor to the assets of the parent, or postponement of the parent's claims, are remedies only. One of the pressing needs in this field is a clear statement of the basis and theory to be used in applying them. General phrases such as "to do equity" are not very exact guides, especially where to do equity to one innocent party necessarily cuts into the equity of other equally innocent parties. Many stockholders in Standard Gas & Electric Co. were quite as innocent as the outside stockholders of its subsidiary, Deep Rock.

17 Taylor v. Standard Gas & Elect. Corp., 306 U. S. 307 (1939).

18 E.g., *In re* Associated Gas & Elect. Corp., 149 F.2d 996 (C. C. A. 2d 1945).

19 321 U. S. 349 (1944).

20 *Ibid.* at 363.

21 In at least one state this result has been reached by statute. Wis. Stat. 1943, § 221.56(3).

22 See Ballantine, Corporations § 122 (2nd edn 1946). The most famous case was an attempt to use the corporate device to thwart the anti-trust laws [Northern Securities Co. v. United States, 113 U. S. 197 (1904)], but the cases are legion.

23 Ballantine, Corporations § 122 (2d ed. 1946); see, e.g., Linn & Lane Timber Co. v. United States, 236 U. S. 574 (1914); Rice v. Sanger Bros., 27 Ariz. 15, 229 Pac. 327 (1924); George v. Rollins, 176 Mich. 144, 142 N. W. 337 (1913); Higgins v. California Petroleum Co., 147 Cal. 363, 81 Pac. 1070 (1905).

24 United States v. Lehigh Valley Ry., 220 U. S. 257 (1911); Chicago Minneapolis & St. Paul Ry. v. Minneapolis Civic Ass'n, 247 U. S. 490 (1917); United States v. Reading Co., 253 U. S. 26 (1919); Brundred v. Rice, 49 Ohio St. 640, 32 N. E. 169 (1892).
25 Brundred v. Rice, 49 Ohio St. 640, 32 N. E. 169 (1892).
26 Northern Securities Co. v. United States, 193 U. S. 197 (1904).
27 E.g., Consolidated Rock Products Co. v. Dubois, 312 U. S. 510 (1941); Southern Pacific Ry. v. Bogert, 250 U. S. 483 (1918).
28 C. C. A. 3d, Aug. 13, 1946. This case is now before the court on petition for rehearing.
29 131 N. J. Eq. 419, 25 A. 2d 418 (Ch. 1942).
30 A striking litigation arising out of such a situation occurs in Blaustein v. Pan American Petroleum Corp., 174 N. Y. Misc. 601 (Sup. Ct. 1940), *modified,* 263 App. Div. 97, 31 N. Y. S.2d 934 (1st Dep't 1941), *modified,* 293 N. Y. 281, 56 N. E.2d 705 (1944).
31 Such an agreement is found, for example, in art. 2, § 6 of the Indenture securing the 30-year debentures of the Mountain States Telephone and Telegraph Co., dated June 1, 1938.
32 This doctrine would lead to a different result in cases where a parent, having operated a subsidiary for years, elects to jettison it without making provision for its creditors or for a fair independent start. In Majestic Co. v. Orpheum Circuit, Inc., [21 F.2d 720 (C. C. A. 8th 1927)] it was fairly plain that the parent, while preserving the separate accounting and operational entity of its subsidiaries, also absorbed their good will and ability to navigate. Yet it was permitted to "dump" one of them, to the loss of its creditors, without liability.
33 Weatherford Ry. v. Granger, 86 Tex. 350, 24 S. W. 795 (1894).

## References


Ballantine, Henry W. (1946), Ballantine on Corporations, 2nd edn, Chicago: Callaghan.

Burdick, Francis M. (1906), "Are Defectively Incorporated Associations Partnerships?" *Columbia Law Review* 6(1): 1–14.

Cook, (1898), A Treatise on the Law of Private Corporations having Capital Stock, 4th edn, Chicago: Callaghan.

Dodd, E. Merrick and Baker, Ralph J. (1940), Cases and Materials on Business Associations: Corporations, vol.1, Chicago: Foundation Press.

Douglas, William O. and Shanks, Carrol M. (1929), "Insulation from Liability through Subsidiary Corporations," *Yale Law Journal*, 39(2): 193–218.

Drucker, Peter (1946), The Concept of the Corporation, New York: John Day Co.

Hornstein, George D. (1941), "Legal Controls for Intracorporate Abuse – Present and Future," *Columbia Law Review*, 41(3): 405–56.

Israels, Carlos L. (1942), "The Implications and Limitations of the 'Deep Rock' Doctrine," *Columbia Law Review*, 42(3): 376–94.

Latty, Elvin R. (1936), Subsidiaries and Affiliated Corporations: A Study in Stockholders' Liablity, Chicago: Foundation Press.

Rembar, Charles (1939), "Claims against Affiliated Companies in Reorganization," *Columbia Law Review*, 39(6): 907–32.

Taylor, Henry O. (1902), A Treatise on the Law of Private Corporations, 5th edn, New York: Banks.

Timberg, Sigmund (1946), "Corporate Fictions: Logical, Social and International Implications," *Columbia Law Review*, 46(4): 533–80.

# 12 The trouble with profit maximization*

*Robert N. Anthony*

*It is too difficult.*
*It is unrealistic.*
*It is immoral.*

"Why are college graduates who are trained in our mother science, economics, so ill equipped to handle real-life business problems?" This question, asked by a businessman some time ago, set me to wondering about a similar one: "Why do graduate students, by applying what they avow are sound analytical tools learned in college, often arrive at naive solutions to the problems in business cases?"

I have finally concluded that the trouble stems from the assumption in most college economics texts and college classrooms that the objective of a business is to maximize profits. Unhappily, this assumption is not confined to the campus. Countless writers of fiction and non-fiction have seemingly taken it for granted that management's purpose is to maximize profits. Lawyers, labor union spokesmen, and government officials often indicate that they share the same belief.

Moreover – as if to confirm that what all these other people think is true – businessmen themselves sometimes *say* they operate on this assumption. One observer, after a study of corporate annual reports, public relations, and other activities, has reported:

> It is surprising and ironical that, to judge by what businessmen often *say*, one would think that they, too, agree that the nature of business corporations is exactly and precisely what critics say it is; namely, that the corporation has no other purpose, and recognizes no other criterion of decision except profits, and that it pursues these profits just as single-mindedly and irresponsibly as it can.[1]

This article is therefore addressed not only to young men in, or preparing to enter, positions of management, but to their seniors as well. And my purpose is not only to help ease the transition from the classroom world of profit maximization to the realities of life but also to remind business leaders of the importance of narrowing the gap between their world as often represented and their world as it actually is.

Before tackling this rather huge task, I should make one point to businessmen readers who may find it difficult to believe that today students are in fact taught profit maximizing economics: they are. I have reviewed what I am told are the five largest selling economics texts, accounting for some 250,000 copies a year.[2] All of them base their analysis of business decisions on profit maximization. One has no qualification – no mention that this assumption may not be valid in practice. Another has a one-sentence qualification. The longest qualification is one of some three pages, but these pages are sandwiched almost exactly in the middle of 251 pages of analysis in which the profit maximization assumption is governing.[3]

Of course, who says what or exactly how much does not really matter. The net effect on students is the only fact that counts, and I can testify to that both from evidence by businessmen and from my own observations of graduate students in the business school classroom.

My argument can be set forth usefully in the form of a catechism.

## Realistic Goals

### QUESTION: *Is profit maximization the dominant objective of American business?*

In general, no. Many companies formed to achieve some specific, short-run objective (e.g., a real estate syndicate, a stock promotion) undoubtedly fit the profit-maximization pattern. So do speculators in both securities and commodities. So do various types of fly-by-night operators and get-rich-quick artists. But I know of no study of general business practice that supports the profit maximization premise, and I shall mention later studies whose findings are inconsistent with it.

Although we find leaders of the business community stressing the importance of a *satisfactory* profit, we also find them discussing business responsibilities, the need to a fair division of income among the parties involved in a business, and other subjects that are incompatible with the profit maximization goal. Donald K. David, for example, has said:

> Business leaders, who wish to preserve and strengthen the kind of society in which we believe, must run a business organization which, beyond being competitive, is a satisfactory social entity for those who work in it and a constructive entity in the national whole. To this basic purpose there must be added the study of the *responsibilities* of management.
>
> The sense of obligation which management must undertake is at least twofold On the one hand, it extends to the people who make up our thousands upon thousands of companies; it means providing for them not only the conditions essential to the effective performance of work but the realizing of their potentialities as persons so that freedom need not be futile or purposeless for any person. The community, which a company itself comprises, must be a *healthy* community, satisfying the noneconomic as well

> as the economic needs of the individuals who make it up and enabling them to consider their work a way of life as well as a livelihood.
>
> But another responsibility, sometimes in apparent conflict with our human commitments inside our companies, extends to the businessman's public responsibilities to the community, to the nation and the world comprised of a such communities.[4]

If profit maximization is the governing objective of business, such a statement is nonsense. And I am quite sure that it is not.

### *QUESTION:* ***Why, then, are economics courses constructed on this premise?***

Probably because its use permits a rigorous intellectual reasoning process. If one assumes profit maximization, a complete and completely consistent package of rules for operating a business can be devised, rules that can be expressed precisely in equations and illustrated by graphs, rules that provide correct answers to classroom problems and rules which, when they do not work in practice, can always be explained by "other things being equal."

The usefulness of such an all-inclusive package for teaching purposes, for the exploration and extension of theories, and as a device for communicating to one's colleagues, should not be minimized. How can one grade an economics examination if there are a whole range of "correct" answers to the problem, or if the correct answer depends on what is "satisfactory" to the individual student? And what does it matter if the assumption underlying the package is not correct, so long as the application of the rules requires rigorous reasoning in solving classroom problems?

In the better economics courses, and especially in advanced courses, students are told that the structure they are studying is a theoretical one, to be explored for its own sake, and not because it conforms to reality.[5] There is a natural temptation, however, not to detract from the interest in the subject by stressing this point, and in the case of many students, this does not sink in.

The professor acts for reasons roughly like those that led tribal medicine men to assume the existence of evil spirits. Medicine men knew of remedies that were effective for certain ailments, but they also knew the great limitations of these simple remedies. By inventing evil spirits who were assumed to cause all illness, and by incorporating the known remedies into a system of potions, incantations, and rituals that were consistent with this assumption, they were able to gain much more influence in the community than if they had relied solely on prescribing the known remedies. Indeed, it is quite likely that the community gained from this, because without this additional influence the medicine men might not have been able to induce their clients to take those medicines that were in fact beneficial.

There is much merit in making apparently reasonable assumptions when the evidence is inconclusive. For instance, there was merit in the notions of the

eighteenth-century physicists about a substance called phlogiston; the assumed existence of which was thought to be necessary as an explanation of the process of combustion. The assumption turned out to be completely wrong, but many believe that its existence facilitated the development of correct principles of physics. Modern physicians, however, have discovered that they can retain the community's respect without the "evil spirit" principle, even though this requires an admission of their inability to cure, or even understand, a great many illnesses.

### QUESTION: *What really is the dominant objective of American business then?*

As a general statement, I suggest that the objective of a business is to use its resources as efficiently as possible in supplying goods and services to its customers and to compensate equitably those who supply these resources. As a way of making this general statement operational, I suggest that the objective be considered as earning a satisfactory return on capital employed (a "satisfactory" return being equitable compensation paid for the use of capital).

### QUESTION: *What is the difference between satisfactory return and profit maximization?*

In the first place, there are a great many problems in economics to which the solution is the same under either assumption These are problems that involve means of achieving goals, rather than the goals themselves. As Herbert A. Simon has pointed out,[6] there are wide differences in the goals of various organizations – businesses, governments, churches, and so forth – but any organization should try to achieve whatever its goal may be in as efficient a manner as it knows how. Marginal analysis, the favorite tool of the economist, is a valid technique for helping to decide which of the alternative solutions to certain problems is the most efficient; that is, which has the lowest cost.

The alternative choice problems where marginal analysis is useful tend to be those where the possible alternatives can be fairly clearly specified, and where it is possible to make reasonably good estimates of the costs and revenue implications of each alternative. Such problems include the acceptance or rejection of certain investment opportunities (Shall we acquire this machine or not?), the choice among various ways of accomplishing a desired result (Shall we buy or lease this machine?), the best way of producing a given mix of products (production scheduling), the best inventory policy for a given sales pattern and set of production facilities, and so on.

Many of these problems are complex and difficult – and many of them involve large sums of money – but in a relative sense they are simple and unimportant; that is, they are simple and unimportant relative to such problems as pricing, choice of the product line, marketing strategy, the direction of research efforts, what size plant to build, and a long list of others. It is problems like these that are not in practice solved by rules based on profit maximization.

Incidentally, relating marginal analysis to the efficiency criterion answers the following question which is embarrassing to the profit maximizers: If marginal analysis depends on the profit maximization assumption, why is the technique so widely used in situations where profit maximization is known *not* to be the objective? The Russians, the nationalized industries in Britain and elsewhere, regulated public utilities, U.S. and other government agencies, and a wide range of nonprofit organizations, all make extensive use of marginal costs for analyzing certain types of problems All these organizations have in common with each other, and with profit-seeking businesses, the task of reaching their objectives as efficiently as possible.

There is another type of situation in which marginal analysis is useful, namely, in times of crisis. In normal times, if a product does not earn a satisfactory return, the businessman is likely to drop it and seek a better replacement unless marketing considerations indicate otherwise. When the going gets rough, on the ether hand, he is apt to keep the product so long as it makes a contribution. He is also more apt to shave prices in bad times – an action based on marginal analysis – although he does not raise prices in boom times to the level that the profit maximization principle would indicate. (The most dramatic example is the automobile companies' decision not to take full advantage of the postwar shortage.)

There are at least two important areas that do involve the objectives of a business, and in which the profit maximization assumption leads to theoretical conclusions that are inconsistent with practice: pricing and capital budgeting.

## What about pricing?

If he is to maximize profits, the businessman must set a price such that the marginal revenue equals the marginal cost. This means that as a minimum he must estimate the demand at all prices and the marginal cost at all volumes, and he must further estimate the extent to which demand is interdependent with cost because of advertising and other order-getting expenditures. In short, he is supposed to estimate all the values for all the lines on Figure 12.1.

This is a fantastically difficult task, so difficult that it is rarely attempted in practice. All studies of actual practice that I am aware of testify to its rarity.[7] Who can accurately estimate the demand for a product at even one price? Instead, the studies show that in most pricing problems the emphasis is on the construction of a "normal" price. Rather than attempting the ritual suggested by the economist's diagram in Figure 12.1, the businessman goes through a much simpler process:

He builds up cost – including direct costs, a fair share of indirect costs, and a satisfactory profit margin; he speculates whether he can probably obtain an adequate volume at a price based on this cost; he considers competitive pressures and strategic matters; and thus he arrives at his price. He may vary the profit margin depending on circumstances, and he may also vary the cost by changing the design of the product, but his starting point is a price based on

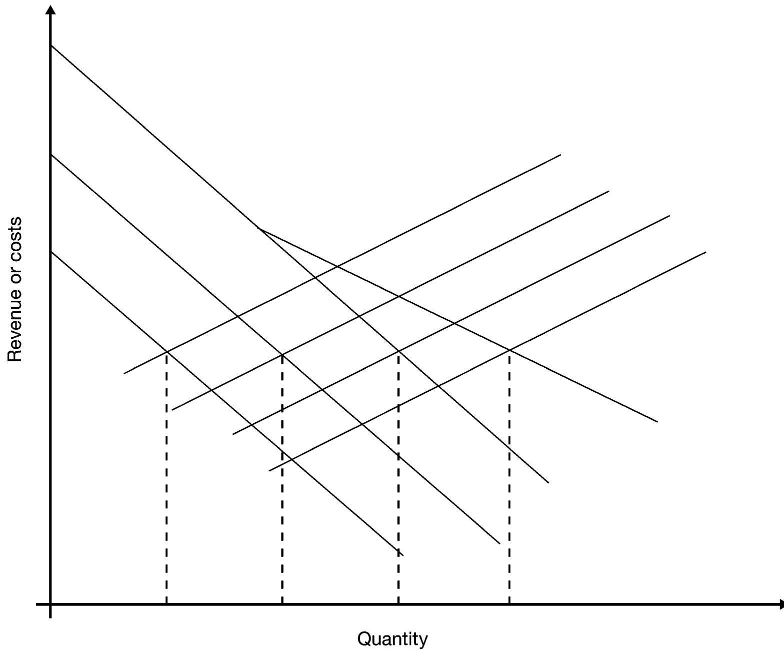


*Figure 12.1* Example of necessary estimates of demand and costs.

total cost as derived from a conventional cost accounting system, not a price based on marginal cost. His reasoning is that if each product contributes a fair share toward overhead costs and profit, then he will make a satisfactory profit on the aggregate of all products.

The profit arrived at by this method will probably not be the maximum profit that could be earned from prices computed from economic analyses, but the businessman is much more comfortable about the likelihood of obtaining this profit than he would be if he relied on the host of guesses required under the marginal approach.

Furthermore, pricing is not the dominant focus of management attention that the economics suggest. Pricing is one element of the total marketing mix which includes also merchandising, branding, channels of distribution, personal selling, advertising, promotions, packaging, display, and servicing.

### *Times for re-examination*

Having established his normal price, he will find that he must re-examine it in many situations. The profit maximizer's approach to these situations, too, is unrealistic. For example, Customer A tells the salesman of a manufacturing firm: "I'll give you an order to Product X if you'll cut 5¢ a pound off the regular price." To decide whether to accept this offer, a profit maximizer requires at least this information:

- The marginal revenue and marginal costs associated with this specific order, which require an analysis of each cost element that may be affected by it.
- The probable consequences of losing the customer's business, both on sales of Product X and on sales of other products, if the order is refused.
- The probable effect on other customers who buy Product X and on buyers of all other products, so as to determine whether idle capacity will exist.
- The probability that the customer means what he says.

Textbook analyses of such a problem usually suggest that the order should be accepted if marginal revenue exceeds marginal costs, if idle capacity is available, and if the impact on future business with Customer A and on other customers is not serious. Some authors even favor a computation of the present value of the expected values of various chains of consequences on prices and costs that might ensue if the order were accepted, or if it were rejected. The only time in which these computations would not be made is when the cost of obtaining the information exceeds the value of the order.

To be sure, there *are* circumstances under which the businessman will undertake parts of such an analysis – when the company is hungry far business, for example. But he will not do this every day for every order. Instead, he will reason in some fashion like this:

> Let's stick to our price. To deviate from it would touch off a lot of reactions, the consequences of which are too difficult to figure. Maybe we will lose some marginal profit, but maybe also we can pick up some unforeseen new business that will produce a profit from the same production facilities.

The profit maximizer tells the businessman to accept the order unless he knows of other orders that will use the same capacity. The businessman ordinarily takes the opposite approach: "I don't have any idea where the orders are coming from, or even if they are coming, but I'm sure going to try to get them." The fact that the sales force will work harder to get business when times are bad is another idea that is not consistent with profit maximization.

It is interesting to speculate what would happen if, when you take your automobile to the garage for servicing, you and the service manager negotiated a price according to the rules of the profit maximizers, with both of you attempting to estimate, among other things, the probability that other cars will use up today's service capacity and the utility of having your car available at a certain time!

If the only relevant costs are marginal costs, then cost accounting with its allocations of depreciation, overhead, and other joint costs to products is a useless ritual, or at most an unnecessarily complicated way of determining inventory values. Some economists indeed assert that cost accounting is useless for pricing. For example take this quotation: "It is difficult to name an occasion when pure cost-plus pricing is in the seller's best interest."[8] Yet practical businessmen would find it very easy to name many such occasions, I believe. And hundreds

of thousands of companies have cost accounting systems that do allocate costs to products, they spend not insignificant amounts to operate such systems, and presumably they think these expenditures are worthwhile.

## What about capital budgeting?

To be consistent with the profit maximization premise, a firm should invest in new assets whenever the return from the investment is equal to or greater than the marginal cost of capital, provided that there is no other available investment opportunity which will permit an even greater return. In theory, therefore, the businessman is supposed always to know his marginal cost of capital, and he is supposed to know about and evaluate all other investment opportunities every time a project is presented for consideration.

In practice, this is too difficult. John R. Meyer and Edwin Kuh, for example, analyzed all the studies of investment behavior they could find, and in only one of them – a 1951 study of public utilities – was there any evidence supporting the validity of the profit maximization premises. They themselves analyzed published data on 750 firms over a five-year period, and found "nothing to justify any claim to unique superiority for any one theory above all other alternatives."[9]

Evidently, businessmen take a much less complicated approach; they set up criteria such as maximum payback or minimum acceptable return, which if things work out as anticipated will ensure a satisfactory profit. This leads to quite different working rules from those prescribed by the economists; indeed, the difference in the literature between articles by economists and articles by practical businessmen on this subject is so great that it is difficult to believe they are writing about the same problem.

## Clarifying the differences

Pricing and capital investment problems are not the only soft spots in the profit maximization concept. Let us proceed with our catechism.

### *QUESTION: Are there other inconsistencies between practice and profit maximization?*

There are a number of them. The practice of corporate donations is one which, although not of great magnitude, is of special interest because the contrast between profit maximization and practice is so clear-cut.

The profit maximizer attempts to explain away a corporate donation to the Community Fund by saying it is made to promote a favorable climate for the company in the community and hence is consistent with long-run profit maximization. But how, then, can he explain the fact that corporate contributions to charity decline in a recession? If the size of the contribution were derived from a maximization kind of reasoning, one would expect either that a recession would have no effect or, since the need is greater and many companies

are quite liquid in a recession, that contributions actually would increase. The obvious explanation is that managements feel *less able* to give when times are bad, a phenomenon non perfectly consistent with the idea that they seek to maintain a satisfactory profit.

Performance appraisal techniques are another illustration. When we are judging how well a manager has performed, the profit maximizer says we must be indifferent to the fact that the business made a profit, or even to the fact that its profit performance is better than that of competitors or better than the results of a year ago. Instead, we should give the management a pat on the back if, *and only if*, we are convinced that it has squeezed out the last possible dollar of profits. Actually, however, it is nonsense to assert that this is the way managements are in fact judged. They are judged, and their bonuses paid, on the basis of improvement or of comparisons against other managements; this is a comparative, not a maximization, idea.

Looking around us at the real world of business, we might raise other questions about the consistency of profit maximization with management practice:

- Why give an employee a separation payment when he is discharged?
- Why don't executives spend all their waking hours at work?
- Why have a lawn around the plant, and why spend money to mow it?
- Why do research expenditures tend to vary directly, rather than inversely, with profits?
- Why are cost-cutting campaigns instituted in times of recession?
- What is the sense of a cost-cutting campaign anyway if costs are always supposed to be at the minimum?
- If prices are always as high as possible, how can a wage or material cost increase lead to a price increase?

### QUESTION: *Are not the phenomena earlier described approximately the same as profit maximization?*

They most certainly are not. A pricing theory that starts with a marginal analysis is completely different from a pricing theory that starts with a full cost analysis. An investment theory built around the cost of capital is in important respects different from an investment theory built around a satisfactory return concept.

### QUESTION: *Are not these phenomena consistent with long-run profit maximization?*

No. It is true that in the long run all costs are incremental, and a pricing calculation made in accordance with a long-run profit maximization premise would contain the same elements as a pricing calculation based on conventional full costs. The numbers in these two calculations would be quite different, however. The profit maximizer's long-run cost calculation must be an estimate

of future replacement costs; the conventional cost calculation deals with current costs. The two would be the same only in an absolute static economy.

***QUESTION: Are business policies and business practices consistent with other maximization concepts?***

No, they are not consistent with welfare maximization – nor satisfaction maximization, nor minimax, nor any of the other elaborate concepts that economists have conjured up in an attempt to find a maximizing theory that can be reconciled with real life. In short, the events we observe in business are not consistent with *any* version of maximization. They are instead related to such notions as "balance," "equity," or "adequacy." The calculus of maximization will not fit these notions any more than it will define what constitutes an excellent dinner, a beautiful woman, a healthy man, a sound tax policy, or an adequate military establishment.

## Impossible ideal

***QUESTION: Should not businessmen maximize, even if they do not in practice?***

It is, of course, appropriate to make a distinction between a descriptive statement and a normative statement, between what businessmen actually do and what they should do. But if the normative statement differs from real life, then the theorist is obligated to explain the discrepancy. Profit maximizers can make this explanation on only two grounds: either businessmen are stupid or they are ignorant. The first is not worth discussing. Ignorance, in its literal sense of "not being aware of," is in fact a valid explanation of why sound new ideas are not immediately and universally adopted. It therefore at least qualifies as a possible explanation.

But the profit maximizer's techniques are not new. Most of them have been around for 50 years or more. Many responsible businessmen were exposed to them as college students. When we observe how such intricate concepts as statistical quality control, economic lot size, linear programming, information processing and other techniques for improving efficiency gain fast and widespread acceptance, while the concepts of marginal income pricing and cost of capital cut offs are not noticeably more prevalent than they were a generation ago – then clearly we have grounds for wondering if the economist's error is perhaps a better explanation than the businessman's ignorance.

There are two excellent reasons why businessmen do not attempt to maximize: it would be too difficult, and it would be immoral. The practical difficulties of applying a profit maximization concept have been discussed in the sections on pricing and capital budgeting. What about the moral aspects?

The ethical problem is that profit maximization requires that the business manager think only of the best interest of the shareholders, whereas any

responsible manager knows that he must actually consider the interest of all parties who have a stake in the business, of which the shareholders are only one. Profit maximization requires the businessman to use every trick he can think of to keep wages and fringe benefits down, to extract the last possible dollar from the consumer, to sell as low quality merchandise as he can legally hoodwink the customer into buying, to use income solely for the benefits of the stockholder, to disclaim any responsibility to the community, to finagle the lowest possible price from his vendors regardless of its effect on them, and so on.

The profit maximizers admit that in doing these things, the businessman must have regard for the long-run consequences of his actions, but as they themselves point out, the "long run" is a long way off and its effect on current decisions is nebulous. They deny the existence of a businessman's conscience, and they exclude ethical considerations as being irrelevant to the subject.

A businessman is a human being, and it is completely unrealistic to assume that he should act in an ethical vacuum. As a human being, he is deeply concerned with how his actions jibe with his own conscience, the respect of his family, and the opinions of his associates. Moral standards change, and whereas 50 or 100 years ago the profit maximizing manager would perhaps have been tolerated in some circles of some communities, today society clearly expects the businessman to act responsibly. He cannot do this and at the same time seek to maximize the share of income going to just one of the several parties that have a stake in the business.

## Social pressures

Laws reflect a society's ethical standards, and by such laws as the Fair Labor Standards Act, Robinson-Patman Act, Sherman Act, Clayton Act, and the many state pricing and labor acts, society is telling business that it disapproves of profit maximization. When President Eisenhower asks the steel companies to hold the fine on prices, he is not displaying the complete naiveté assumed by the economists; rather, he expresses the nation's conviction as to how steel companies should act. It is blindness to build a theory on the premise that businessmen completely disregard such beliefs. In fact, I doubt very much if an economist could imagine himself running a business with no ethical standard, as required by profit maximization.

In short, businessmen could not maximize if they wanted to, and they would not want to if they could.

In his normative statements, the economist assumes that the businessman should be an "economic man" – an omniscient, completely rational, unfeeling, amoral automaton. The social psychologists have long since given up theories based on the thesis that the individual worker is an "economic man," nor are theories of the market based on the assumption that the consumer is an "economic man" regarded as having any validity today. Why do the profit maximizers believe that Adam Smith was right about managers when they know that he was wrong about workers and consumers?

## The alternative

### *QUESTION: Does not the satisfactory return idea lead to a loose set of concepts?*

The concepts of a satisfactory return model are indeed not as precise as those erected on a profit maximization base. However, to criticize the satisfactory return assumption on this account is comparable to criticizing the physicists for their acceptance of Heisenberg's uncertainty principle. In both cases, the resulting body of theory is less precise, but it is also more realistic.

The satisfactory return assumption does not imply that decision rules are swept away and that management actions become entirely subjective. The satisfactory return for a particular business can be described – not precisely, but within reasonable limits. The lower limit is the company's expected cost of capital, and the upper limit is related to the profit opportunities inherent in the industry. Within these limits, the figure will vary with circumstances – chiefly the aggressiveness of the management and its attitude toward risk and growth.

Nor does it follow that the acceptance of this assumption means that income is distributed according to the dictates of management, as implied by Mason.[10] Labor unions, boards of directors, investors, bankers, and the government all exert pressure to ensure that the group each represents receives an equitable share of the revenue. Mason is quite right in pointing out that the share going to each group cannot be precisely determined by the satisfactory return theory, whereas the profit maximization theory does give a precise answer to this problem. The trouble is that this precise answer is wrong.

It is equally fallacious to argue that the satisfactory return concept implies that managers are lazy. Certainly they will vigorously seek out opportunities to improve profits when they can do so ethically, and competition will force them to seek ways to improve efficiency even if no increase in profits results. In our vigorous, dynamic society, considerable effort is required merely to hold one's own.

Let me suggest that those who believe that satisfactory return is not an adequate stimulus to management ought to take a try at managing before they publicize such opinions.

## Improvements needed

Admittedly, the concepts based on the satisfactory return idea are far from adequately worked out at present. This is because economists have only recently begun to work on them. If, for example, economists interested in price theory turned their attention away from the fine points of marginal income analysis to the problems of constructing the most useful full cost, they could make important contributions. And, incidentally, such a change in focus would have the happy consequence that economists would become less patronizing toward cost accountants. In the capital budgeting area, they could develop useful approaches to the many perplexing problems that arise in attempting to select the most

attractive investment opportunities from those available. They might even make important contributions to accounting theory, which is now a curious mixture of concepts implying profit maximization and concepts implying satisfactory return.

The model of a business that results from all this work will be less exact than the model that the student now learns about. This inexactness should not be glossed over; it should instead be stressed in much the same way that models of weather systems or of community relationships are taught as being only rough approximations of reality. Much less attention should be given to the fine points implied in the over-all model and much more attention to an analysis of the realities of specific business situations. The latter is more valuable than drawing diagrams that require omniscience and clairvoyance to duplicate in practice.

## Conclusion

I have tried to show that profit maximization is not a valid assumption to explain either how businessmen actually behave or how they should behave. I believe that this assumption is unrealistic because (i) profit maximization is too difficult and (ii) it is immoral.

The consequences of the profit maximizers' misinterpretation are not only that their concepts are not useful to businessmen, but also that they happen to be conveying a false impression of what our economy really is like. Consider this statement, which is part of Samuelson's summary of the essence of our economic system:

> A rich man's dog may receive the milk that a poor child needs to avoid rickets. Why? Because supply and demand ere working badly? No. Because they are doing what they are designed to do – putting goods in the hands of those who can pay the most, who have the money votes.[11]

This is a shocking statement. If it were true, no one should be proud of the American system; such a system would certainly not be welcomed in other countries. If it were true, we should prefer communism.

The plain fact is that this statement is not true. Our system does not condone, let alone encourage, fattening dogs by starving children. Our system *is* one of which we may be proud. It can be described accurately if the assumption of profit maximization is discarded for the idea of satisfactory return. Such a change will lead to note accurate reporting and to the development of more useful rules and concepts, concepts which focus on the businessman's responsibility to all the parties at interest, concepts which we can be proud of, and which will lead to improvements in our system of which we can be even more proud.

> And let [the merchant] know that disappointment comes easily with too much greed in seeking advantage, and excessive pursuit of gain is the road

to loss. The explanation consists in this: that between the purchase of one who has a wild desire to buy and the purchase of another who has a faint desire and who heals his soul of the madness of greed and keeps it free from the slavery of passion, there is a wide chasm and a great difference.

Abu al-Fadl Ja 'far ibn 'Ali al-Dimishqi,
"The Beauties of Commerce" (Damascus, late ninth century),
in *Medieval Trade in the Mediterranean World*, edited by
Robert S Lopez and Irving W. Raymond.
New York: Columbia University Press, 1955.

## Acknowledgment

We hope this reprinted contribution may eventually become a tribute to Prof. Robert N. Anthony who sadly died on December 1, 2006.

## Notes

* Adapted from *Harvard Business Review*, 38 (6), November–December 1960, 127–34.
1 Glover (1954: 328).
2 Bach (1957); Dodd and Hasek (1957); Harris (1959); Homan, Hart and Sametz (1958); Samuelson (1958).
3 Samuelson (1958: 181–3).
4 Address before the Transportation Association of America, January 20, 1954.
5 A number of authors who criticize the validity of the profit maximization premise are referred to in the excellent article by Julius Margolis (Margolis 1958). See also March and Simon (1958). Another exceptions is Edgar M. Hoover (1954); but note how the conventional economists castigated him (*American Economic Review*, May, Suppl, 228–35). See also Edward S. Mason (1958).
6 H.A. Simon (1947).
7 One of the earliest of these studies was reported in "Price Theory and Business Behavior" (Hall and Hitch 1939); see also Robert F. Lanzillotti (1958).
8 Alfred R. Oxenfeldt (1958).
9 Meyer and Kuh (1957: 204).
10 Mason (1958).
11 Samuelson (1958: 12).

## References

Bach, G.L. (1957), *Economics: AnIntroduction to Analysis and Politics*, Englewood Cliffs: Prentice-Hall, Inc, 2nd edition.

Dodd, J.H. and Hasek, C.W. (1957), *Economics, Principles and Applications*, Cincinnati: South-Western Publishing Co., 3rd edition.

Glover, J.D. (1954) *The Attack on Big Business*, Boston: Division of Research, Harvard Business School.

Hall, R.L. and Hitch, C.J. (1939), "Price Theory and Business Behavior," *Oxford Economic Papers*, No 2, Oxford: Clarendon Press.

Harris C.I. (1959), *The American Economy; Principles, Practices, and Policies*, Homewood: Richard D. Irwin, Inc., 3rd edition.

Homan, P.T., Hart A.G. and Sametz, A.W. (1958), *The Economic Order; an Introduction to Theory and Policy*, New York: Harcourt, Brace and Company.

Hoover, E.M. (1954), "Some Institutional Factors in Business Investment Decisions," *American Economic Review*, May, Suppl, 201–13.

Lanzillotti, R.F. (1958), "Pricing Objectives in Large Companies," *American Economic Review*, December, 921–40.

March J.G. and Simon H.A. (1958), *Organizations*, New York: John Wiley & Sons, Inc.

Margolis, J. (1958), "The Analysis of the Firm: Rationalism, Conventionalism, and Behaviorism," *Journal of Business*, July, 187–99.

Mason, E.S. (1958), "The Apologetics of Managerialism," *The Journal of Business*, January, I.

Meyer, J.R. and Kuh, E. (1957), *The Investment Decision, An Empirical Study*, Cambridge: Harvard University Press.

Oxenfeldt, A.R: (1958), *Pricing New Products*, AMA Management Report No. 8.

Samuelson, P.A. (1958) *Economics, An Introductory Analysis,* New York: McGraw-Hill Book Company, Inc., 4th edition.

Simon, H.A. (1947) Administrative Behavior; A Study of Decision-Making Processes in Administrative Organization, New York: The Macmillan Company.

# 13 The entrepreneur: the firm*

*James H. Stauss*†

The term "entrepreneur" as ordinarily employed in economic literature is a holdover from the classical era of political economy, when individual proprietorships and partnerships were predominant and when those who possessed the ownership equities in corporations were supposed to exercise effective decision-making power. Now, in its customary definitions, it fits neither the requirements of modern theory nor the historical pattern marked by concentrations of power in government, labor groups, and officers of the firm. A revision of the concept is appropriate. Recent discussion has emphasized the importance of this fact; current usage of the term is unsettled, although its definition and the considerations underlying choice of a definition are vital both to an accurate account of the structure and functioning of an economy and to research in economic history.[1]

Accordingly, this paper is an argument to advance the proposition that the *firm* is the *entrepreneur.* In order to support this proposition, traditional definitions must be examined and the reasons for rejecting them made clear. Furthermore, the nature of the firm must be explained, the significance of its decision-making organization emphasized, and the roles of those functionaries having relations with the firm orientated. The ensuing analysis is in consideration of these requirements. With respect to the proposition that the firm is the entrepreneur, it must be set forth at the beginning that the entity known subsequently as *the firm* is taken as a real institution. As such the firm *exists* apart from the individuals who compose its decision-making organization, but it *does not function* apart from them. Thus the entity is not a fiction; it is a fact.[2]

## Section I

Proceeding with the intention of delimiting a certain function or group of functions undertaken by a class of individuals (i.e., by the so-called "entrepreneurs"), economists are not in agreement as to the unique role associated therewith. In general, the familiar definitions are either historical (pertaining to an individual typical of a given time and place) or formal (pertaining to an individual undertaking a specified responsibility through the medium of the firm, regardless of the structure of the economy). The historical definition, culminating in the notion of the *captain of industry,*[3] evidently is suited to a typical phenomenon of business but one that has come to have limited historical significance with the recent development of corporative firms, both private and public. In this sense, the entrepreneur is a character of economic history. The formal definition, designed to specify a peculiar agent of production, is in consideration of one or

the other, or both, of two fundamental responsibilities: the assumption of risk and the assumption of management. In this sense, the entrepreneur is a character of economic theory. It is with this definition that the subsequent comments are concerned.

The English classical economists and Marshall, though not making the convenient distinction between historical and formal definition, regarded the entrepreneur as the individual who, as primary financier of the firm (typically, an individual proprietorship), must choose the opportunity for employment of resources and for establishment of the creditor equity (in event of credit financing). Moreover, and basically, they regarded him as the *proprietary capitalist manager*, jointly the risk-bearer and (at least) the decision-maker, whether alone in the firm or with others of his kind.[4] But their discussions of risk were deficient in analysis of the concept itself[5] and were not integrated to make evident the fact that the assumption of risk is a responsibility undertaken generally by all supply functionaries in the employment of economic resources and in the possession of equities. Their discussions of management were also defective, owing to the absence of a thorough treatment of the function of control, especially in regard to its relation to supervision, coordination, and decision-making in view of uncertainty. The entrepreneur was supposed to exercise control because of the power derived from his position in the ownership of the firm. Consequently, all possessors of the ownership equity must come within the scope of the entrepreneurial concept, which can be generally valid only if all members participate in the basic responsibilities believed to substantiate it. However, in view of this requirement, a hiatus is evident in the classical definition of the entrepreneur, for the possessor of a share of the ownership equity, especially under the joint-stock principle though exercising the responsibility of assumption of risk, may have a very tenuous relationship to the managerial function.

Whatever the reasons for change in definitions, the main current of thought bifurcated. Say, Walker, J. B. Clark, and Schumpeter were instrumental in emphasizing and refining definitions of the entrepreneur based upon the managerial function.[6] On the other hand, Hawley and Knight worked toward establishing the coherency underlying risk-bearing and management – in effect, to close the classical hiatus by relating risk-bearing to control.[7] There are, however, cogent reasons for questioning the adequacy of these concepts.

Relative to the various definitions, each of which asserts a unique entrepreneurial role, it is essential to test the validity of segregating in one class the so-called "entrepreneurs." Is the supposition correct that these functionaries assume certain responsibilities either primarily or exclusively? In order to facilitate this purpose, the classical definition must be subjected to examination. Since the revisions of the classical definition are different only in refinement of the fundamental elements of the latter, they may be analyzed concurrently.

Although it is the property of the firm that is immediately subject to risk in consequence of the firm's obligations, the possessors, or the possessor, of the ownership equity are often considered to be specialists in risk-bearing. For, while

the risks of other groups are more remote, theirs is primary in regard to their equity and income and may be in regard to their property and income not directly involved in the firm. Thus defined, the entrepreneurial role reflects the responsibility of *provisional guaranty* of the equities, resources, and contractual remunerations of others. Any risk which includes this responsibility is a primary risk. However, the qualifications that need to be made are sufficient to dissolve the notion of a universally primary risk borne uniquely by the specified group. These are the qualifications.

1 In the event of high principles of ethics or trust, private or governmental, or in the event of suitable organization, perhaps forceful, the responsibility of guaranty might not be required.
2 The possessors, or the possessor, of the ownership equity might share this responsibility with underwriters of some kind, or the latter might assume complete responsibility.
3 The risk assumed by the possessors, or the possessor, of the ownership equity is not primary to many of the risks assumed by other groups; and, when it is primary, owing to the possibilities of inadequacy and subterfuge, the adjective "provisional" may, on occasion, not mean very much.
4 The ownership equity is often characterized by gradations of responsibility for the guaranty, with certain grades more remote than others, and the creditor equity, also often characterized by similar gradations, may differ from the former as a matter of degree. This is illustrated by the range of equities in the corporation.
5 A further complication of the whole subject develops when the responsibility is a governmental burden. In view of these qualifications, the conclusion must be that the concept of a primary risk is not a unified one and that other groups may share the responsibility of provisionally guaranteeing certain obligations of the firm despite the more remote positions of their risks. That the risk-bearing of the possessors, or the possessor, of the ownership equity cannot be recognized as of a unique class is reinforced by:
6 The observation that the risk associated with the creditor equity in one firm may be preferred to the risk associated with the ownership equity in another firm.

It appears, then, that if risk-bearing reflecting the responsibility of provisional guaranty must be considered an integral part of the entrepreneurial role, despite the inability to associate it only with the risk-bearing of the possessors, or the possessor, of the ownership equity, two alternatives are present: (i) to broaden the scope of functionaries undertaking the entrepreneurial role or (ii) to save the classical definition by modifying it in accordance with other doctrines purporting to meet the objections outlined. The first alternative must be abandoned, for that would result in ambiguous identification of the functionaries. The flow of productive services, relative to technical constituents of the plant, depends, except as limited by existing obligations, upon the willingness (not necessarily

voluntary) of certain ownership groups, that is, of laborers (considered here as owning themselves), owners of capital resources, and possessors of equities. These agents, who assume the responsibility of relations with the firm to the end of implementing the supplies of factors of production, may be designated as "supply functionaries"; they are not factors of production. The assumption of risk, basic to their ownership status, may involve numerous types of liability to loss, among the significant sources of which is *the choice of one employment over alternatives*. It is tautological, and indeed misleading, to assert that these agents are also entrepreneurs to the extent that they are risk-bearers. Risk-bearing is a substantial element of the role of the supply functionary, whether laborer, owner of capital resources, or possessor of an equity share. Consequently, the second alternative presents itself. But an evaluation must await the subsequent discussion of the managerial function, because certain arguments of significance there are intended to bolster the classical definition.

Tentatively, the managerial function may be considered as involving supervision, co-ordination, and decision-making in view of uncertainty, although these components are difficult to separate.[8] The classical definition makes the possessors, or the possessor, of the ownership equity primary in undertaking these responsibilities, although a reasonable interpretation of this view would seem to limit the function in many cases to a final, effective control. Nevertheless, as pointed out before, these functionaries often appear to be exempted from the control function, notably in many instances of the modern corporation. When that is the circumstance, the problem arises whether such individuals are entrepreneurs under the classical definition.

The symmetry of the classical definition may be said to have been established by Knight, who has defined the function of ultimate control (or ultimate management) as that of selecting men to make the decisions required in the conduct of the firm.[9] It is a function granted to those exercising it because they are provisional guarantors of the equities, resources, and contractual remunerations of others.[10] Accordingly, the exemption of various possessors of the ownership equity from the managerial function would be held to be fallacious, for they would participate in the selection of hired executives, at least through exercise of legal voice in the affairs of the firm (as in the election of directors) or through tacit submission to the selection made by others (as indicated by the decision to possess the equity share and to perform no office).

Knight, however, in not limiting the function of undertaking provisional guaranty to the possessors, or the possessor, of the ownership equity (and in some cases excluding them altogether), has set up a broader class of entrepreneurs: the responsible directors, or the provisional guarantors, whoever they may be, that exercise the function of ultimate control. Thus, the class of entrepreneurs would include either (i) those who give under contract a provisional guaranty of the equities, resources, and remunerations of others or (ii) those who participate in the selection of hired executives through an effective consultatory jurisdiction or only through the tacit submission involved in the decision to possess an equity or to employ resources.[11]

The English classical economists did not attempt to separate labor activity and entrepreneurial activity. In associating the entrepreneurial role with the function of responsible direction, Knight has emphasized such a separation.[12] On the other hand, Say, Walker, Clark, and Schumpeter have defined the entrepreneurial role in terms of particular types of entrepreneurial labor. The last, as noted before, has asserted the proposition that the fundamental feature of management is the introduction of new combinations of productive factors. Accordingly, the crucial decision is supposed to be the decision to effect such introduction.

Consequently, it appears that if management reflecting the responsibility of making crucial decisions, whether by the so-called "responsible directors" or by other significant functionaries, must be considered an integral part of the entrepreneurial role, several alternatives merit examination: (i) to save the classical definition (as supplemented by Knight) by viewing it as setting forth a standard type, though subject to deviations for which account is necessary; (ii) to delimit the scope of functionaries undertaking the entrepreneurial role, in order to conform to one of the revised classical definitions; or (iii) to propose a new concept of the entrepreneur that will place in proper perspective the functions associated with both risk-bearing and management. The subsequent argument follows the third alternative, in conformity with the opinion that, although Knight's solution of the problem of control is logically correct, it is neither the only solution nor one that is sufficiently relevant by the criterion of correspondence to the data of experience.

## Section II

Under the traditional concepts of the entrepreneur, that functionary was regarded as exercising his responsibilities through the medium of the firm. Thus, the firm would lose much of its significance as a functional entity. The frame of reference in treating of the vast subject of enterprise would be a class of agents considered to have primacy either in meeting the risks involved in production or in undertaking in some manner the crucial decision-making which rules production, or else primacy in these activities jointly. On the other hand, if the frame of reference is the firm, an entity operating through the medium of various functionaries having heterogeneous responsibilities and relationships to it, the exposition would be consistent with the significant facts of experience, which are contrary to the concept of a unique class having primacy in undertaking the function of risk-bearing or that of management. It would also be consistent with the point of view that the entrepreneurial role should be described in a manner to locate, relative to specific problems, the effective authorities in management and to integrate in relation to these authorities the structure for meeting risk. Thereby the treatment of enterprise would take on a realism, symmetry, and clarity not found when the main introduction to the subject is by way of a class of individual functionaries. The basic proposition of this revision is that *the firm* is *the entrepreneur* and that the central relationships of entrepreneurship are those of decision-making by the firm.

A fundamental fact of enterprise is that decision-making in the firm conditions risk, since the liability is directly affected by errors in managerial judgment. The decision-making organization of the firm, which assumes active risk-taking, and the judgments which it renders are basic factors shaping both the risk-bearing of functionaries having contractual relations with the firm and the risk-bearing of the firm itself. Liability to the latter as a functional entity is implied because it possesses assets and receives income. But the cruciality of certain decisions is not something to be determined on abstract grounds by means of an analysis presuming the universal pre-eminence of limited hypotheses; it is a matter that depends upon the circumstances and problems of the time and place. The focus of analysis cannot be diverted from decision-making that occurs within the jurisdictional spheres of the respective functionaries employed in the firm, where plans are devised and administered, where new combinations are introduced, where much liability to error is incurred, and thus where risks are directly and significantly conditioned. In no empirically important sense are these activities routine. Furthermore, the nonroutine, crucial significance of many types of control decisions must be established. Since the classical definition (as supplemented by Knight) requires criticism from this point of view, it may now be examined.

In the modern economy the structure of the firm is highly complex. Two developments are basic: (i) the corporation in which the roles of ownership and effective decision-making are largely separated, in view of the decentralization of private possession of the ownership equity,[13] and (ii) the corporation in which the roles of ownership and effective decision-making may be separated, in view of private ownership and governmental regulation, or are unified, in view of governmental ownership, regulation, and administration. In regard to the first development, it appears perfunctory to assign the possessors of the ownership equity or the so-called "provisional guarantors" (largely the same group) primacy in exercising control on the ground that they govern the selection of hired executives, at least through exertion of legal voice in the affairs of the firm or through tacit submission to the selection made by others. Substantially, such participation may have no appreciable influence on either the continuous employment of executives or the decision-making which occurs within the jurisdictional spheres of the various employees. In selecting the board of directors, several important types of control have been distinguished:[14]

1 control by means of centralization of the ownership equity in possession of an individual or small group of associated individuals;
2 control by means of centralization of a majority of the ownership equity carrying voting rights;
3 control, without centralization of a majority of the ownership equity carrying voting rights, by means of a legal device (as pyramiding, a class of the ownership equity carrying special voting rights, or a voting trust) designed to assemble power to select the board of directors;
4 control by means of centralization of a minority of the ownership equity carrying voting rights, implemented by executive officers willing to select a

favorable proxy committee and by the ability to secure enough proxies to cast a majority of the votes at the periodic election; and
5 control by executive officers through possession of the proxy machinery, sometimes reinforced through the use of legal devices and other methods. There are further types of control, too, e.g., through organization of interests associated with the creditor equity, through activity of labor unions, through banking consent not to withhold financing, and through administration by governmental agencies. Decisions thus made govern the selection of men and determine the bounds of jurisdictional spheres within the firm.

Primacy in exercising the function of control cannot be located formally in any one class of so-called "entrepreneurs" grouped on the basis of some other uniform relation to the firm (as possession of the ownership equity or extension of provisional guaranty). The circumstances of time and place demand consideration in locating the controllers and deciding upon the importance of their decisions. Likewise, decision-making within jurisdictional spheres is in the face of managerial problems of greater or less importance concerning the conduct of the firm. Cruciality of decisions is thus a relative matter. Frequently it is not fruitful to attempt the location of powers of control as a general phenomenon, for in the varied activities and requirements of the firm primacy in exercising the function may not exist continuously but may be distinct for each of many situations involving conflict of interests. Therefore, in view of the evidence considered, the conclusion is that the possessors of the ownership equity are not of a unique class having primacy in the decision-making of the firm. Furthermore, the necessity of emphasizing the heterogeneity of functionaries and of methods in exercising control, as well as the necessity of stressing the agents and activities in the jurisdictional fields of management, if the theoretical introduction to enterprise is to be complete, suggest that the central relationship of entrepreneurship should be the decision-making of the firm. When the firm is treated as the entrepreneur, a comprehensive frame of reference is at hand to be utilized in examining the details of particular cases and in formulating typical instances in regard to the functionaries and methods involved in the exercise of the function of management.

When the second corporate development, reflecting the tremendous influence of the governmental role on enterprise in peace and in war, is considered, the conclusions derived before are reinforced. Even though the corporation be under private ownership, governmental decisions and pressures, made through commissions, other governmental agencies, and various authorities, may dominate the policies of the firm in many respects, possibly to the extent of governing the selection of hired executives. Ownership and control are amalgamated in the public corporation; but this unification is unrecognized in terms of the responsibilities of the functionaries supposed by traditional definitions to have primacy in effective decision-making.[15] The latter depends upon many variables, including legal and extra-legal voice of the citizen or the group in affairs of the state, machinery and operation of politics, governmental custom, and personal

influence of officials. In view of the heterogeneous institutions of government, no unique class of functionaries, through translation of its political powers into powers relative to the affairs of the firm, is primary in exercise of the function of control in the public corporation. Its exercise is, at most, a matter of typical instances and, at least, a matter of distinct authority in dealing with particular situations. There is no uniform pattern of effective decision-making.

To recapitulate previous conclusions in respect to the assumption of risk and the assumption of management, the facts of experience suggest that the responsibility of risk-bearing generally and of provisional guaranteeing particularly are not undertaken uniquely by the possessors, or the possessor, of the ownership equity. Furthermore, they suggest that the exercise of the function of control does not reside in a unique entrepreneurial role associated therewith. Therefore, the classical definition depends for its validity upon a clear recognition that it is representative only of a standard type: the individual possessor of the ownership equity or of a share in it undertaking through the medium of the firm the responsibilities of provisional guaranty (relative to risk-bearing) and of control (relative to management), regardless of the extent to which either is exercised under practical conditions.

However, this definition, setting forth the entrepreneur as a standard type, is unsuitable for the purpose of theoretical analysis. For it does not conform to either of these criteria: (i) If a functionary performing a particular role is a standard type capable of conversion into a fundamental category of theory, that functionary must have such dominant pragmatic importance that exceptional agents in undertaking the same responsibilities are neither numerous nor significant. *Or*, (ii) if the standard type is valid, it must have such antecedent importance logically that it is a first approximation to the description of exceptional cases, the essential characteristics of which are most precisely interpreted as contrasts to those of the standard. As the preceding exposition has indicated, the possessors, or the possessor, of the ownership equity is not the only important agent engaged in either risk-bearing or control of the activities of the firm. Furthermore, since the essential characteristics of the possessor of the ownership equity or of a share in it are, apart from certain technical details, identical with those of the exceptional functionaries (i.e., with the other functionaries also marked as risk-bearers or provisional guarantors and as crucial decision-makers), the former is not of antecedent importance logically; all functionaries so related depend for their common ground on the same characteristics. The description of one is not a first approximation to the description of the remainder.

It is not only, however, because the standard type itself is incapable of conversion into a fundamental category of theory that the classical definition of the entrepreneur is not suitable for the purpose of theoretical analysis. The irresistible fact is that an adequate presentation of economic principles requires a *general* concept which effectively integrates the whole exposition of enterprise – a concept with ramifications that embrace the scope of risk and management in production. Traditionally, the analysis of enterprise has been approached through consideration of a group of individuals, i.e., of so-called “entrepreneurs,” who

assume certain responsibilities through the medium of the firm. This latter entity, though implicitly integrating the structure for undertaking the entrepreneurial role, has been, *per se*, secondary. The point of view must be reversed, however. The firm is the entrepreneur. The firm, centered around its decision-making organization, operates as the functionary in undertaking the entrepreneurial role through the medium of individuals having relations with it. That such a concept of the entrepreneur is logically adequate is the fundamental proposition of this paper.

## Section III

Turning now to the definition of the entrepreneur proposed by Schumpeter and to the different definition proposed by Knight,[16] two questions must be posed: (i) Is the specified function (respectively, the introduction of new combinations or the ultimate selection of men to make the decisions required in the conduct of the firm) pre-eminent in a general survey of the workings of the economy? (ii) Is the suggested class of functionaries identified unambiguously, in so far as an introductory pattern of the structure of the economy is concerned?

As pointed out previously, Knight's definition is inadequate, because the nonroutine, crucial significance of other types of control and decisions is not recognized. On the other hand, Schumpeter's definition rests on the separability of entrepreneurial and other activity, in accordance with the proposition that a distinctly economic change, *per se*, is involved in the choice and introduction of new combinations and in the attitude of thought which nourishes such choice. However, enterprise (decision-making and consequent activity by the firm, or risk-taking – in contrast to risk-bearing) is more than innovation; it is also dealing with the variance between the outcome of provision for a contingency and the anticipated outcome in making the provision. For example, the firm must deal with frequent employment crises (relative to men and materials) that arise in the course of adjustment to powers of organized groups and government. But, whatever might be said about the proper scope of activity involved in the process of initiative, Schumpeter's analysis of discontinuous innovation is readily fitted into the *economic* framework viewed here as the assembly of firms, supply functionaries, consumers, and government. Thus, the functionaries introducing new combinations would be, in the main, firms (old or new) acting through their aggregates of individual members with specified powers of decision. Consequently, the identity of laborers in the role of factors of production or in the role of supply functionaries would be preserved.

With the firm considered as the entrepreneur, the introduction to enterprise would depend upon a frame of reference competent to emphasize the basic responsibilities of laborers to yield productive services and to initiate and continue relations with the firm. Such are the pre-eminent functions of laborers in the initial broad view of enterprise, in the sense that the first order of observation is to recognize those functions and functionaries responsible for the continuation of production and for the use of productive services. It is true that the exposition

of economic principles must be more fruitful when the labor category is dissected and when those functionaries who are the bearers of economic change are segregated under some nomenclature. But such dissection belongs to a subsequent order of observation.

It is ambiguous to assert that specified laborers or specified members of ownership groups are also entrepreneurs to the extent that they undertake the particular function believed to signify the entrepreneurial role. The rendering of managerial services is the substantial element of the role of certain laborers as factors of production, or else (when decision-making does not involve labor activity) is most accurately considered as incidental to or pertaining to the fundamental role of certain supply functionaries – laborers (considered here as owning themselves), owners of capital resources, and possessors of equities. The main function of these latter agents, which have relations with the firm, is implementation of the supplies of factors of production. A secondary function is involved when participation in selecting hired executives is exercised by those who give under contract a provisional guaranty or by those who have an effective consultatory jurisdiction or who tacitly submit to outside selection by the decision to possess an equity or employ resources. To identify these supply functionaries also as entrepreneurs is simply to employ a second term to designate a category already familiar; and the duplication of terminology is just as misleading as when managerial laborers, familiar as factors of production, are in addition designated as entrepreneurs.

Nevertheless, in connection with Knight's definition of the entrepreneur as the responsible director, the question must be propounded whether those who give under contract a provisional guaranty do not occupy a special status as supply functionaries – a status differentiating them from other supply functionaries owing to the crucial significance of their decision-making or risk-bearing. Before this question is answered, it is necessary to delimit more precisely the members of the class of provisional guarantors of the equities, resources, and contractual remunerations of others. One means of distinguishing this group of so-called "entrepreneurs" is to define them as recipients of a *profit income*, thereby to emphasize a functional service supposed to be performed by its recipients.[17] Thus, a profit income is taken to be any functional income (to laborers, owners of capital resources, or possessors of equities) which is dependent upon the prior satisfaction of all functional incomes which are *promised*. It is a residual income because it is contingent upon meeting legally recognized promises of payment to specified claimants and is itself unpromised. The recipients of such an uncertain share are deemed to be unique because only through their willingness to accede to these terms of remuneration can production be carried on. Theirs is a primary risk-bearing, and it carries with it the assumption of responsible direction of the affairs of the firm.[18]

Such a location of the entrepreneurial class is inadequate because, despite the fact that its members may have to demonstrate their willingness to have production carried on, the essence of decision-making in the firm does not reside universally in them and also because willingness as risk-bearers to implement production is not their own province exclusively.

The continuity of production does not result only from the willingness of recipients of an unpromised, residual income to accede to such terms of remuneration. It also results from the willingness of functionaries whose remunerations are promised to persist in their various relations with the firm. Consequently, both classes undertake the function of control, if assumption of the function is evidenced by the decision not to cease their respective activities. Although the selection of decision-making personnel for the firm and the determination of jurisdictional spheres within the firm may be accomplished directly through choice by specified groups with active power, the decision of various functionaries to continue relations with the firm is fundamental; it constitutes, from the formal point of view, final confirmation of the selection of men, the determination of jurisdictional spheres, and the continuity of production. However, such confirmation is made by agents in their capacities as supply functionaries, whether they be laborers, owners of capital resources, or possessors of equities. The recipients of unpromised income are not unique as functionaries upon whose willingness the continuity of production depends.

No violence is done to this conclusion by recognition of those past decisions whereby contractual obligations once assumed may be instrumental in providing for current willingness to continue relations with the firm, whether by the recipient of unpromised income or by any other supply functionary. The decision by an original contractor is not necessarily an act stipulating that this identical party retain his initial relations with the firm but may put him in a position of having to substitute another functionary should he terminate his interest. For example, the original possessor of an equity share, either as a stockholder or as a bondholder, made provision for future willingness to fulfill the responsibility of such a functionary, although he might transfer the responsibility by sale of the certificate. Of course, to the extent that contractual obligations assumed in the past are binding, current willingness to continue relations with the firm is not the result of completely uncontrolled decisions. Thus, looking to the past, the organization and continuity of the firm may have required that original contractors other than the recipients of unpromised income show willingness that enterprise be initiated and carried on.

When the decision of functionaries to continue relations with the firm is considered to be the main frame of reference in locating the source of control, the solution is not sufficiently relevant to the data of experience, especially in view of the two corporate developments emphasized heretofore. The cruciality of too many other decisions is unrecognized, despite the fundamental importance of the activity upon which it is based. Another solution provides a frame of reference which is more adequate and formally logical: to make the firm the entrepreneur and to locate the exercise of the function of control in the decision-making organization of the firm, whatever its structural characteristics in a particular case. Such a revision gives a comprehensive scope to the function of control, because it may be utilized to take account of the multiplicity of authorities responsible for active control, whether relating to selection of personnel, to consultation or negotiation in matters involving private interest

groups, to administration by governmental agencies, or to jurisdiction within managerial spheres. As maintained before, analysis of control in terms that associate primacy with a more passive role appears perfunctory, affording the exposition of principles a limited and scarcely realistic retrospect of the workings of enterprise in the economy. If the interpretation of decision-making in the firm is to be comprehensive, the various activities involving management must be integrated, with active control and all crucial decisions not secondary and derivative but at least complementary to the willingness that production be continued. Therefore, the decision-making organization of the firm must be taken as the amalgam of functionaries carrying out these activities of management and, consequently, as the factor of unification.

If the concept of the entrepreneur as the recipient of unpromised income is unsatisfactory in respect to association of that agent with the function of control, it remains to decide whether the concept is adequate when such a functionary is considered to be one who assumes a peculiar risk – peculiar because there exists no contractual promise of income (and, in some cases, no contractual promise of repayment of investment principal). However, the risks of promisors and promisees are different only as a matter of degree, in so far as priority in allocation of the firm's gross income or in liquidation of equities is concerned. Of course, some risks are marked by the absence of a promise of remuneration and are thus set apart as unique in kind, although this characteristic is often provisional, depending upon the absence of specified circumstances under which the legal rights of the risk-bearer to remuneration would become effective. Nevertheless, this peculiarity in kind is a *technical* fact; it is aside from the *economic* fact of risk as a factor in the supply of economic resources. The willingness of laborers, owners of capital resources, and possessors of equities, whether promisors or promisees, to continue relations with the firm involves the assumption of risk by these agents. They are risk-bearers in their capacities as supply functionaries, and as such their particular risks differ in many technical details, as illustrated by stipulations relative to promise of remuneration, ownership of the firm,[19] underwriter's guaranty, or rights of control. Therefore, if the entrepreneurial category is to continue to serve as a basis for the introduction of economic principles in their relation to enterprise, the structure of risk-bearing cannot be divided to distinguish the promisor and the promisee for the purpose of designating the former as the entrepreneur.

## Section IV

An analysis of enterprise proceeding from traditional concepts of the entrepreneur must be incomplete: emphasis on the assumption of risk or on the assumption of management, to the exclusion of the other, in distinguishing the entrepreneurial role results in omission of an ingredient of enterprise as well as in failure to coordinate the two; and emphasis on the assumption of responsible direction neglects the function of control and the rendering of crucial decisions in so far as power to act is not immediately derived from the provisional guaranty under

contract of the incomes, equities, or rented resources of other groups. Many of the difficulties involved in the problem of defining the entrepreneur can be cleared up, however, by recognizing the many elements in the role of supply functionaries. The bearing of risk, involving willingness of functionaries to initiate and continue relations with the firm, is a responsibility united with the function of control in so far as the latter also involves the same willingness to engage in economic activity. Nevertheless, in order to give complete form to the discussion of entrepreneurship, the inseparability of active risk-taking and managerial decision-making must be stressed. In conformity with previous argument, it is clear that risk-bearing must be merged with risk-taking, because the state of exposure to a specified loss is fundamentally conditioned by decision-making in the firm. In turn, the firm, through its decision-making organization as an aggregate, whatever the groups of which that aggregate is composed, undertakes this active risk-taking, and the firm as a functional entity is the antecedent risk-bearer, because it possesses assets and receives income. Thus, a unique, integrating agency itself provides a frame of reference for a comprehensive analysis of enterprise and of the relations of the various functionaries to the organization of production.

This revised orientation would place the risk-bearing functionaries in their proper positions as agents having relations with the firm – as laborers, owners of capital resources, and possessors of equities but not as entrepreneurs so defined because of some factor associated with these capacities. The firm would stand out as a separate entity engaged in dealing with supply functionaries, in employing factors of production, and in disposing of the commodity thereby obtained. Its prime mover would be the decision-making organization, composed of an aggregate of individual members with specified powers of decision due to participation in management, either as laborers or as supply functionaries with powers of control.

At this point, however, the problem of properly defining the firm arises. In accordance with previous argument, two alternatives present themselves: (i) to define the firm as an *association* or aggregate of the members of the decision-making organization or (ii) to define the firm as an abstract entity, for which the members of the decision-making organization make decisions and with which supply functionaries have specified relations.

If the first alternative rules, the point of view may be that the firm is different whenever the decision-making organization changes.[20] Or, on the other hand, the point of view may be that the firm has continuity of existence, on the grounds that many interpretations of phenomena require the concept of an institution having the qualities of a *going concern*.[21] The authority of Commons, who pioneered economic thinking along the latter line, probably supports the associational concept of the firm,[22] although, without distinguishing it, an entity concept is also recognized.[23]

The main criticism that must be urged against this definition of the firm as an association or aggregate of certain members is that the concept is a general and integrating one because it is so inclusive. The concept is not one of a unique,

integrating functionary that itself provides a frame of reference for a comprehensive analysis of enterprise and of the relations of various other functionaries to the organization of production. Unification of the entrepreneurial role would not be paralleled by unification of the structure by which it is undertaken. Instead of an entity with primacy in the exercise of decision-making (involving risk-taking), with other groups having relations with such an entity, there would be many separate entities to be accounted for in location of the entrepreneur. In fact, an account of the entrepreneur would be a general description of all functionaries engaged in carrying on production. Thus, the employment of a fundamental category of theory (the entrepreneur), in order to give introduction and perspective to the substantial details of economic life, would be abandoned.

Commons' concept of the going concern, when applied to the aggregate of members of a decision-making organization (defined as the firm), differs from that of an association of individuals because the composite is personalized or considered to have a collective will. In turn, the collective will is defined as the *working rules* relative to which the members function.[24] However, the same criticism must be leveled at this view. The firm would be located and its structure would be described only by taking account of many diverse functionaries or, alternatively, by taking account of many diverse rules and customs. An exposition designed to delineate certain significant relations as carried on by some principal has placed upon it too large a burden when the description of that principal becomes too detailed and complex for formal treatment. Furthermore, the notion of *personality* apart from human beings is bound to encounter strong opposition from many sources.

Therefore, the second alternative (that of defining the firm as an entity) must rule, if the logic underlying its choice is correct. Prior to considering the character of this entity, it is necessary to set up a model as the principal assumed to be endowed with the capacity of undertaking certain relations (i.e., decision-making in the sense of risk-taking). The concept of the firm would be analytically unmanageable were the firm considered as an association of decision-makers, because, as pointed out, there must be a lack of regularity among diverse rules and among members viewed primarily in their capacities other than as components of a decision-making organization. Thus, the abstract principal is representative of the association in so far as the framework for carrying out the entrepreneurial role is itself significant. The device is in lieu of the aggregate which is too cumbersome to handle. Nevertheless, despite its nature as an abstraction, the expediency of employing such a model in scientific work is well recognized.[25]

Now this model taken as a copy of the decision-making organization may be institutionalized (i.e., may be clothed with the character of an economic mode of thought and action) in contemplation of economic theory and practice. For the notion of a unit confronted by income calculations over time is indispensable to modern theory.[26] And the firm is regarded, in a pragmatic sense, as a reality, intangible as to body, but understandable as a unit with which transactions are

carried on and by which factors of production are employed. It is often so understood by those who have market, as well as legislative, administrative, and judicial, relations with it. For certain purposes in accounting it is treated as a unit having a separate existence and financial condition.[27] And for certain purposes in the law it is treated as a legal entity with its rights and obligations.[28]

As an abstract entity (a model functionary, representative of the decision-making organization and recognized as an institution), the firm may be defined as an *accounting entity*, subject to possession of assets and subject to calculation of its revenue and other income, costs of revenue and other costs and allowances, and income and other financial allocations. Within its limits as an accounting entity, it is a bargaining and transacting agent, dealing with supply functionaries, consumers, and government on terms that permit of quantitative calculation and estimate. Accordingly, the firm (the entrepreneur) is a fundamental category of economic theory serviceable as *one* frame of reference in a preliminary view of the economy in the light of basic functions and classes of functionaries. In its *substantial* analysis relative to a particular problem and its data, it is peculiarly institutional and must be made concrete in consideration of the details of structures for making decisions, implementing production, and employing resources.

With the firm recognized as the entrepreneur, another problem presents itself: to construct conceptual standards as frames of reference in analyzing the *income to the firm* or to reconstruct the theory of profits as income to the entrepreneur. A primary factor in an interpretation of revised concepts of the entrepreneur and profit must be *uncertainty*, an essential accompaniment of historical change. In view of uncertainty, anticipations govern the decisions of both supply functionaries and the decision-making organization of the firm. When profit is defined as a residual after deducting "necessary costs of production," this fact may not be sufficiently taken into account. "Necessary costs" are frequently viewed as remunerations required in the "long-run" in order to evoke the economic resources whose services are essential to the continuity of production. And "going rates" of remuneration are then taken as criteria of "necessary costs" to be deducted from revenue in arriving at profit as a residual income. In contrast, however, are the facts of historical change: change in "going rates," although with temporary inflexibility in many rates of remuneration (owing to the terms of contract or law); anticipations of what "going rates" may rule at future dates; and frequent changes in such anticipations. Decisions are made under these conditions. Evidently economic theory lacks a standard by which to define "necessary costs of production" in the historical "long-run."

Profit concepts apply to incomes estimated as realized in the past, in relation to calculations that are records, and they may also be made to apply to incomes estimated as possible in the future, in relation to calculations that are anticipations under alternative plans. In turn, the plans of the decision-making organization, formed as they are in consideration of the data of experience and historical insight, will usually reflect the objective of improving in some manner, for the benefit of some interest, the *firm' s position* as a residual claimant. Therefore,

anticipations of profit (defined according to some accounting standard) and techniques for quantitative arrangement of these anticipations are fundamental matters in analyzing the entrepreneurial role.

## Notes

* Adapted from *Journal of Political Economy*, 52 (2), June 1944, 112–27.

† James H. Stauss (Iowa 1914 – December 19, 1975) took a BA at Grinnell College in 1936, followed by a MS (Economics) at Iowa State University in 1937, with a master's thesis on "The Economics of Horace Greeley," and finally a Ph.D. at University of Wisconsin, in 1945, with a thesis on "The Entrepreneur: The Firm" that he synthesizes in the article reprinted here (biographic note by Yuri Biondi).

1 Arthur H. Cole, "Entrepreneurship as an Area of Research" in *The Tasks of Economic History* (supplement to the *Journal of Economic History*, December, 1942, 118–26; Frank H. Knight (1942); J.M. Clark (1942); George Herbert Evans Jr (1942).

2 Institutions, like the Congress, the university, and the church, are frequently understood as noumena, in a manner substantially the same as that proposed relative to the firm. People are cognizant of their existence as entities apart from the individuals who compose their functional organizations at a given date.

3 Cf. Helen Wright (1930): 216–20; Werner Sombart (1930): 406.

4 Alfred Marshall (1920): 293–304, 357–8, 589–91, 596–603,612–23, 745–7; Mill (1909): 405–7. I have used the word "proprietary" to refer to the characteristic of possession of title to the commodity resulting from employment of productive resources. Many economists consider this characteristic as an attribute of the entrepreneur as an individual; but such a point of view, while consistent with the older assumption that the firm is a collection of individual investors dominated by the owners, is contrary to the newer assumption that the firm is an entity distinct from the possessors of equities in it (cf. Paton and Littleton 1940: 7–11).

5 I shall here use the term "risk" in an eclectic sense: the liability (or state of exposure) to a specified loss, the contingency of the loss being measurable or unmeasurable. This provides a comprehensive frame of reference for analyzing a particular case involving an uncertain event and associated contingencies as to its occurrence, error of judgment, and resulting losses.

6 The proprietary manager: J.-B. Say (1853): 79–85, 296–7,314–18,329-32. The proprietary organizer, planner, and chief executive of industrial operations: Francis A. Walker (1887): 60–1, 73–7, 173–4, 188, 232–45, 279–82,348,387–8. The proprietary co-ordinator and supreme decision-maker under "dynamic conditions": John B. Clark (1907): 83–90, 117–24, 155–8. The functionary introducing new combinations of productive factors: Joseph A. Schumpeter (1934): 19–22, 61 n., 65–94, 102, 128–56.

7 The proprietary (or, more generally, factor-owning) risk-taker: Frederick B. Hawley (1907): 9–15, 57–67, 90–101, 106–59, 175, 214–15, 257–9, 272-81, 306–51. The responsible guarantor of contractual obligations and selector of managerial decision-makers: Frank H. Knight (1921); "London School Reprints of Scarce Works," No. 16: 233–312.

8 Cf. Nicholas Kaldor (1934): 67–70; Austin Robinson (1934): 248–55; P. Sargant Florence (1934): 726–9; George J. Stigler (1941): 385–7.

9 Knight (1921): 291–302.

10 Knight (1921): 270, 277–8, 299, 355–6.

11 Knight (1921): 300–2, 304, 349–50. If the class of entrepreneurs is broadened to conform to the second interpretation, the relation between responsibility and ultimate control or direction becomes thinner, to the point at which responsibility involves risk-bearing generally, rather than provisional guaranty under contract. Knight's exposition

admits of either interpretation, because it is not clear as to the exact nature of the responsibility required before responsible direction, rather than direction alone, is present.

12 According to Knight's theory, the primary function in management is the selection of men to render the decisions required in the conduct of the firm, and the primary structure of the decision-making organization is a hierarchy of functionaries in which those of each level, relative to that below, select men and evaluate their abilities, and, relative to that above, are selected and have their abilities evaluated. Therefore, each functionary leaves the consequences of his activity to his selector; responsibility is continually shifted to a higher level, until it finally rests with the controlling functionaries or functionary, the guarantors, or the guarantor, of the equities, resources, and contractual remunerations of other groups, who ultimately decide on the selection of personnel and thus direct the firm. The latter decision is crucial, and the activities of all subordinates, whether in decision-making concerned with the selection of men or in decision-making concerned with problems of general co-ordination within the respective jurisdictional limits, are of a routine nature (cf. Knight (1921): 267–70, 276–7, 291–302, 309). However, an incidental, and contrary, view of decision-making is also stated (Knight (1921): 358–9).

13 Cf. Berle and Means (1932): 1–68.

14 Berle and Means (1932): 69–90.

15 There has been some recognition that the entrepreneurial role need not be undertaken by individuals (cf., e.g., Walker (1887): 482–3; Maurice Dobb (1932): 558–60). Furthermore, Marshall apparently was on the verge of associating the entrepreneurial role with the decision-making activity of the firm (considered *per se*, a functional entity) (Marshall (1920): 138–9).

16 Cf. above, nn. 6 and 7. Schumpeter's definition of the entrepreneur is a direct result of his view of the mechanism of internal economic development (i.e., of change as a phenomenon of discontinuous introduction of fresh combinations of productive services, involving practical innovations in techniques, organization, and markets). This internal development is in contrast to that change which involves, per unit of time, a continuous adaptation to increments or decrements of data. Furthermore, the theory is that such development or occurrence, abstracted from historically causative factors which change the data – the wants, resources, and technology – is purely economic. Although economic life is determined by historical factors (regardless of what they are or how they are interpreted), the power of the determinants is transmitted by a process and attitude of initiative that is fundamental *per se* and is itself change on a different theoretical level. The entrepreneurs are those through whose initiative the new combinations are effected (cf. Schumpeter 1934: 57–66, 79–94, and Schumpeter 1935: 2–10).

17 Cf. Ben W. Lewis (1937): 535–44.

18 Lewis (1937): 537–41.

19 Cf. Charles A. Tuttle (1927):13–25, for an illustration of a definition based upon this technical detail.

20 Cf. Kaldor (1934): 67–70. The criterion of change in the decision-making organization may be hard to establish. If it is a change in the amount of co-ordinating ability, the standards of measurement are undeveloped. Nevertheless, aside from the problem of defining the firm, such a criterion may be essential far certain purposes.

21 Cf. John R. Commons (1924): 143–213; Paton and Littleton (1940): 9–11.

22 Commons (1924): 144–6.

23 Commons (1924): 147–53.

24 Commons (1924): 147–8.

25 Cf. Cohen and Nagel (1934): 367–75.

26 Cf., e.g., Robert Triffin (1940): 93–95; George J. Stigler (1942): 147–86; Kenneth E. Boulding (1942).

27 Cf. William A. Paton (1922): 472–8.
28 Cf. Edward H. Warren (1929): 1–33, 134–40, 268–74, 291–2, 293–301, 319–23, 418, 425, 841–6.

## References

Berle, Adolf A. and Means, Gardiner C. (1932), *The Modern Corporation and Private Property*, New York: Harcourt, Brace & World.

Boulding, Kenneth E. (1942), "The Theory of the Firm in the Last Ten Years," *American Economic Review*, 32 (4): 791–802.

Clark, John B. (1907), *Essentials of Economic Theory*, New York: Macmillan.

Clark, John M. (1942), "Relations of History and Theory," *Journal of Economic History*, 2 (S): 132–42.

Cohen, Morris R. and Nagel, Ernest (1934), *An Introduction to Logic and Scientific Method*, New York: Harcourt, Brace & World.

Cole, Arthur H. (1942), "Entrepreneurship as an Area of Research," *Journal of Economic History*, 2 (S): 118–26.

Commons, John R. (1924), *Legal Foundations of Capitalism*, New York: Macmillan.

Dobb, Maurice (1932), "Entrepreneur," in Edwin R.A. Seligman and Alvin Johnson (eds), *Encyclopaedia of the Social Sciences*, vol. 5, New York: Macmillan.

Evans, George H., Jr. (1942), "A Theory of Entrepreneurship," *Journal of Economic History*, 2 (S): 142–6.

Florence, P. Sargent (1934), "The Problem of Management and the Size of Firms: A Reply," *Economic Journal*, 44 (176): 723–9.

Hawley, Frederick B. (1907), *Enterprise and the Productive Process*, New York: Putnam.

Kaldor, Nicolas (1934), "The Equilibrium of the Firm," *Economic Journal*, 44 (173): 60–76.

Knight, Frank H. (1921), *Risk, Uncertainty and Profit*, Boston: Houghton Mifflin Co.

—— (1942), "Profit and Entrepreneurial Functions," *Journal of Economic History*, 2 (S): 126–32.

Lewis, Ben W. (1937), "The Corporate Entrepreneur," *Quarterly Journal of Economics*, 51 (3): 535–44.

Marshall, Alfred (1920), *Principles of Political Economy*, 8th edn, London: Macmillan.

Mill, John Stuart (1909), *Principles of Political Economy with some of their Applications to Social Philosophy*, edited by W.J. Ashley, London: Longmans, Green & Co.

Paton, William A. (1922), *Accounting Theory, With Special Reference to the Corporate Enterprise*, New York: Ronald Press Co.

Paton, William A. and Littleton, A. Charles (1940), *An Introduction to Corporate Accounting Standards*, Chicago, IL: American Accounting Association.

Robinson, Austin (1934), "The Problem of Management and the Size of Firms," *Economic Journal*, 44 (174): 242–57.

Say, Jean-Baptiste (1853), *A Treatise on Political Economy*, trans. C.R. Princep, 4th edn, Philadelphia, PA: Grigg & Elliot.

Schumpeter, Joseph A. (1934), *The Theory of Economic Development: An Inquiry into Profits, Capital, Credit, Interest and the Business Cycle*, trans. Redvers Opie, Cambridge, MA: Harvard University Press.

—— (1935), "The Analysis of Economic Change," *Review of Economic Statistics*, 17 (4): 2–10.

Sombart, Werner (1930), "Capitalism," in Edwin R.A. Seligman and Alvin Johnson (eds), *Encyclopaedia of the Social Sciences*, vol. 3, New York: Macmillan.

Stigler, George J. (1941), *Production and Distribution Theories: The Formative Period*, New York: Macmillan.

—— (1942), *The Theory of Competitive Price*, New York: Macmillan.

Triffin, Robert (1940), *Monopolistic Competition and General Equilibrium Theory*, Cambridge: Harvard University Press.

Tuttle, Charles A. (1927), "The Function of the Entrepreneur," *American Economic Review*, 17 (1): 13–25.

Walker, Francis A. (1887), *Political Economy*, 3rd edn, New York: Henry Holt & Co.

Warren, Edward H. (1929), *Corporate Advantages without Incorporation*, New York: Baker, Voorhis & Co.

Wright, Helen (1930), "Captain of Industry," in Edwin R.A. Seligman and Alvin Johnson (eds), *Encyclopaedia of the Social Sciences*, vol. 3, New York: Macmillan.

# Part IV

# Essays on economic, legal and accounting features of the firm as an entity

# 14 Accounting and the economic nature of the firm as an entity*

*Yuri Biondi*

## Introduction

This chapter deals with the accounting system as a "clue" for understanding the economic nature of the firm as a whole and a dynamic system. The dynamic entity view supported by accounting has two fundamental implications. First, it implies increased recognition of the special economic and monetary process generated by the whole firm that is autonomous from external markets. Second, it enhances the economic understanding of the separation between ownership, control and management beyond the irrevocably lost proprietary sovereignty. This leads to an interdisciplinary approach linking economics, accounting, and law by the shared, synthetic notion of the firm as an entity.

The following text is organised in four sections. The first section summarises the three fundamental steps required for the development of a new theory of the nature of the firm integrating the dynamic accounting view: the critiques of equilibrium firm and of incomplete-contracts firm, and the understanding of the firm as an entity, a whole, a dynamic system.

The following sections further explain these steps. The second section concerns the innovative but bounded contribution of incomplete-contracts economics, which neglects both the logical and functional implications of the *firm-entity,* while the third section, starting from some genuine insights of Coase, Shubik, and Simon, but enhancing them in accordance with the recent developments of Baker, Gibbons and Murphy and Rajan and Zingales, explores the fundamental role of the accounting system in *constituting* this *whole.* The last section further develops such a theory of the economic nature of the firm-entity as a whole jointly constituted by management, organisation, the accounting system, business incomes and results, and the institutional environment.

## Looking for the theory of the firm as an entity

### *Equilibrium theory of the firm* as *a theory of the firm,* but *a misleading one*

Equilibrium economics constitutes the usual framework for theories of the firm. The box that equilibrium theory designs for the firm appears neither *black* nor

*empty*. This *marginal cost pricing* box seeks to be a theory of the single firm, that is, both to grasp the underlying economic and monetary *process*, at least in its ultimate elements and results, and to explain selling price, cost, quantity, and resultant (null) profit for each product separately.[1] Also called the "production function," this skeletal machinery, together with some "principle of maximisation" and with perfect competition on all the related markets for factors or products, depicts the firm in a fundamentally static fashion. Firm *dynamics* is relegated to alleged exercises, and the equilibrium hypothesis is always maintained. This bundle of instruments *allows the price system alone to dominate the firm,* at least from the economic viewpoint, when creation and allocation of resources are concerned. Economic systems other than prices in equilibrium do not exist.

In fact, like in other domains, the Emperor equilibrium has no clothes (Kirman 1987). Concerning the firm, the kingdom "that has to be investigated is so vast [that] as we discover it, we can better see the limits of a purely market theory of the firm" (Perroux 1966: 16).[2] Instead of equilibrium theory, a novel perspective[3] appears to be required to understand the firm as an institution and an economic organisation. Framing the firm's economic *activity* with equilibrium imposes some *essentially static* coherence between the framework and the actual behaviour of the firm. Yet experience teaches us also that business economic activity is subject to other modes of existence, for contingencies and uncertainties are no less given in its special activity than are regularities and order. Furthermore, actual resultants of behaviour vary between firms and in real dynamics, subject to different patterns of management and organisation, and cannot provide a general basis for theorising.[4] Thus, the fully predictable pattern derived from equilibrium theory has to be, to say the least, incomplete. As *functional* frame of reference for economic *behaviour*, equilibrium theory of the firm consists of a straightforward, over-simplified connection between what that can be, what that is or will be, and what that ought to be.

Consider the metaphor of a seed and of its potential tree. According to equilibrium theory, under certain determining conditions, the firm as a seed could become an actual tree, and *nothing else*. Human entities, and the related social and economic dynamics, are not like the atomistic seed. They not only deal with the natural world but also with the human world, made up of liberty and responsibilities. They are human activities ultimately concerned with purposes and fairness, with what shall be done right or wrong. In the case of the seed, even though discrepancies between facts and theory occur, the equilibrium "seedology" seems to tell us enough about general principles and actual experience. But the case of the firm goes against the seed. The economic analysis framed by equilibrium represents at best an *ex ante* mode of functioning based on atomistic elements or individuals alone. It neglects the genuine implications of the actual complexity of the whole firm confronted with real time and uncertainties.

> Time and uncertainty have essentially disappeared from this apotheosis of the price system. But it is time and uncertainty which are the concerns of

> everyday economic life and the problems of how to account for the influence of time and uncertainty in the ongoing economic process are central to the development of accounting.
>
> (Shubik 1993: 228)

In conclusion, the resilient, enduring activity of the *firm-entity* as economic organisation and institution relates to its going, stable existence in real dynamics (going concern), but also to the evolution of competencies and resources as we experience in its creative or destructive development (becoming concern). Understanding the interactions between general purposes (the ends) and the means at disposal, as well as between potential aims and actual results, begs indeed for a different frame of reference and analysis.

### *The innovative but bounded contribution of incomplete-contracts economics*

In fact, a different framework is provided by incomplete-contracts economics (both new institutional economics, and property rights theory), but it is none the less constrained in searching for a deeper understanding of the business economic activity called the firm.[5]

Under the *price system*, institutions are relegated to an exogenous framework that does not play any active role in the economic and monetary process. Entities are neutral (just "black boxes") and institutions simply do not matter. In a like manner, accounting also has been neglected for a long while, relegated to the margins of this process. As all other institutions, accounting did not matter. The price system was the sole comprehensive mode of regulating, organising, and knowing the firm.

The "new institutional revolution" fostered by Coase deals with the opposite insight: institutions matter and the firm deserves its special economics, alternative and complementary to the *price system*. Nevertheless, incomplete-contracts economics does not deal very well with important legal and accounting features of the "structure of production" implied by the firm:

- Its notion of "property rights" is not what the law and legal principles hold and what judges have been applying in the context of corporations;
- Its notions of assets, liabilities, costs, and revenues do not fit the accounting views, principles and norms as currently applied by firms and economic organisations.

Its view of the "institutional structure of production"[6] and the related notion of *property rights* do not fit the law, economics, and accounting of the firm. Furthermore, legal structure is limited to individualistic property rights (however defined), neglecting other varieties of legally-enforced norms dealing with, for example, accounting-based constraints. Finally, incomplete-contracts economics holds a narrow *contractarian* view focusing only the firm as a *nexus of contracts* and

of ill-defined property rights. However original and fruitful its contribution may be, its framework stands solely on individual bargaining and suffers from the surviving influence of the myth of the *lonely entrepreneur*, a proprietor and equity provider managing his own business.

Incomplete-contracts economics, therefore, underdevelops some genuine insights of Coase himself, its early promoter. Concerned with the allocation of resources, Coase had especially stressed: (a) the *economic* distinction between firm and market, i.e. how the firm *supersedes* the price system; (b) the active role of legal institutions in the special economic process hence generated by the firm; (c) the inner working of organisation and of accounting systems as key tools for this special economics of the firm.

> In this planned society, the firm costs do not, in the main, arise directly out of the operations of the market but are computed and provided by the accounting system. While outside the firm prices and therefore costs are explicit (because of the demands of others for resources) and are determined by the operations of the market, within the firm there are explicit costs *for exactly the same reason* but they are provided by the accounting system. This internal system takes the place of the pricing system of the market.
>
> (Coase 1990: 11, italics added)

The danger with such a *dualism* between the firm and the market is to misunderstand the effective interaction and the nature of both,[7] for example by personifying the firm, or the market. Dualists see every economic interaction as *make or buy* and *own or hire* decisions. This comes to a two-terms system, or dialectics, between two merely contrasting (or complementary) terms (see Figure 14.1).

From the economic viewpoint, the market is not a person and, in particular, cannot act as a person, nor can the firm. More properly, the market is a *price system* which these new theories call upon in order to interact with the firm, as an alternative but complementary mode of functioning of economic activity. The firm as a whole does not act, but constitutes the *managed economic system* that creates and maintains the favourable conditions allowing the business activity to become and fulfil, if possible.

## *Outlines for a further theory of the firm as an entity*

A new framework of reference and analysis is required to grasp the inner working of the firm and its active role in the economic and monetary process, while dealing with individuals and other entities.

Firm ———————————————— Market

*Figure 14.1* The usual dualism between the firm and the market.

In the "black box," systems other than prices in equilibrium do not exist: we thus have the price system and the firm as a nexus of prices framed by some "principle of maximisation" and by paramount efficient markets (price system). In order to overcome this bias, understanding the firm as a *nexus of contracts*, even though incomplete, is not enough. In fact, the firm is not only a legal fiction which serves as nexus for a series of contracts among individuals or proprietors, devoted to the quick pursuit of immediate shareholders' wealth: neither legal nor accounting logic and principles currently share this view on the firm and its role in economy and society.

Therefore, the emphasis is on the need for a new perspective, one which understands the firm as a managed and organised *system of relationships* not only of the contractual or bargaining sort. The firm provides the special field (the entity) in which individuals and structures mutually interact. In this special economic environment, both principles and choices, purposes and constraints, order and disorder, efficiency and waste, honesty and guile, development or distress have much to do with *structures* of such relations, more than existing theories have already recognised.

In analysing the firm, this approach pays careful attention to three different "structures" (that we would better name "systems") of internal organisation, related: (i) to the nature of learning, diffusion, and information processes (epistemic dynamics); (ii) to the closeness of products, technologies, resources and internal organisation (organisational dynamics); and (iii) to the framing of working rules and norms dealing also with financial matters and regulation (institutional dynamics). This dynamic holistic view allows a better understanding of the legal and accounting features specific to the inner working of the ongoing firm.

A new interdisciplinary approach is hence developed, linking economics, accounting and law by means of a unique synthetic notion: *the firm as an entity*, still common to the three fields and understood here as a *whole* (according to institutional law and economics) and a *dynamic system* (according to accounting and continental business economics, in turn related also to old institutionalism).

Looking for this new transactional and institutional perspective, the dualism between the "price system" and the "firm as economic organisation," however sophisticated, is insufficient to understand their mutual interaction and, more importantly, the nature of the firm. In contrast to the received dualist position, we suggest a *five*-terms system. Five *constituents* are selected to analyse the firm's becoming economic activity under real dynamics and complexity. The three inner constituents are: (a) management, (b) organisation, and (c) its special process. The two outer constituents are: (d) the incomes and results generated through from this process, and (e) all the undertakers (stakeholders) looking for them, related, more generally speaking, to the institutional environment and to the human community (the entity's world). This five-terms system would be better suited to understand the economic nature of the firm as an entity.

Concerned with this ambitious purpose, the article outlines an alternative frame of reference and analysis, starting from the *accounting system* as viable

alternative to *marginal cost pricing* in understanding the economic and monetary process as actually emerges and evolves into the whole firm.

In a world of real complexity (moreover, combined with real dynamics), the "black box" – i.e., the straightforward logical chain linking such atomistic elements as selling price, quantity and cost for each product separately – is disregarded as ultimate frame of analysis for the entity's economic *process*.[8]

Dissecting a whole into separate elements yields not the *constituting* terms of the whole that is analysed, but just a number of new parts. When an element becomes part of a whole, it ceases to be a separate unit, as the whole is not merely assembled with disparate units, but appears as a *constituted* enduring pattern, existing and functioning as such. In this context, the interaction of the parts is not sufficient to understand the enduring existence and functioning of the firm as a whole. Under real dynamics and complexity, the firm-entity further appears as a special business *activity*, analysable in terms of modes of functioning (constituents) which jointly constitute its actual process of becoming a whole.[9] Each mode is analysable within the entity as an autonomous *active* component, and "in its *interaction* with" ("playing an *active role* in") the *economic and monetary process* with which this special activity is concerned. This process is therefore "constituted" or "structured" in the sense that it arises from the features of the working structures of the parts' relationships as much as from the individual attributes (or behaviour) of the parts actually involved.

In this perspective, the economic nature of the firm is grounded on the fundamental relationship between the *actual economic coordination* provided by (a) the firm as a whole, as a *managed* system, as an entity, and (b) the inner organisation of (c) its *going economic process*. This coordination is especially seen as a potential generator of (d) emerging business incomes and results for (e) all the undertakers (stakeholders).

## The nature of the firm as economic organisation: the entity's missing connection

Both the question of the nature of the firm, and the related distinction between the firm and the market, are fundamental issues for the new theories of the firm. Three major approaches to these questions are: agency (AT) and property rights (PRT) theories on the one hand, and transaction costs theory (TCT) on the other. The first two neglect any difference of nature between the firm and the market. The firm is seen as a legal fiction the reality of which is the simultaneous *nexus of contracts* between the individuals engaged in business. These theories stress the *ex ante* incentive structure framed by strategic equilibrium bargaining with incomplete contracts. Provided specific investments exist, the approach would relate only to the asset(s) owned by each proprietor separately.

On the contrary, TCT provides a more sophisticated view, dressing this methodological contractualism in a transactional and institutional fashion. Because of market failures, asset specificities have to be protected by an alternative institutional structure: the hierarchy. Therefore, the original notion of

values *specific to the transactions* could refer to the *inter-individual* nature of this structure, building on the ultimate constituents that jointly allow the firm as a whole to exist beyond the contracting parts.

Unfortunately, this kind of logical and functional development is neglected. According to Williamson (1988), all three approaches share the same perspective (incomplete-contracts economics), based on (i) *opportunism* and *moral hazard* implied by individuals' motives and behaviour, shaped by (ii) the bargaining nature of their contractual interactions, and seeking economic efficiency by means of (iii) endogenous *ex ante* institutional rules (PRT) or *ex post* governance structure (TCT), such as the "Board of Directors."

Even though the notion of entity exists in accounting theory and regulation, this framework completely overlooks the firm-entity as a whole because of its compromise with equilibrium theory, and of the primary connection between the firm and its proprietors, i.e., the providers of equity finances. But the entity phantom actually appears in each approach:

- AT usually *personifies* the firm-entity as a specific actor, either the proprietary-manager or the manager-agent, and makes him pay the agency costs which are the basis of its analytical machinery;
- The notion of *property rights* provided by PRT does not understand assets and claims as the law (Hodgson 2002; Kirat and Bazzoli 1998), economics (Hölmstrom 1999, Demsetz 1995), and accounting (Scott 1979) represent them. For example, this notion ignores the collective nature of property within the firm, which allows the entity to own and possess the assets.[10]
- Finally, TCT does not take into account the economic and dynamic implications for off-contractual matters of its notion of hierarchy as economic organisation. How can the firm and the market remain symmetric in this approach? Given asset specificities, the market can never replicate the firm. Thus, *undertakers* of this kind of investment have to be protected otherwise. In this manner, the firm-entity as a specific institutional environment relates to the economic interdependencies between undertakers (transactions), and can properly justify the existence of values *specific* to these.[11]

Here all the new theories of the firm are maintaining several ties to the quite archaic character of the *lonely entrepreneur*, proprietor and equity provider, managing his own business.[12] This capitalistic hero takes and bears the risks and endorses the management (supervision, coordination, decision-making, control) of its economic activity. In this context, the firm is relegated to a legal and economic device for entrepreneurial action. Decision-making and control relate essentially to the *ownership* of the firm, and the functional modes of existence as a whole disappear.

This surviving idea conceals fundamental facts concerning the firm and prevents effective development of many genuine suggestions provided by a transactional and institutional perspective.

Law and accounting tell us about the functional distinction between firm-entity and owners: the entity as a whole owns and possesses the assets, is able to

assume its own obligations and has priority rights in collecting economic and monetary streams and results. Law and regulation first protect actors other than owners. Ownership consists of certain subordinated rights to ultimate liquidation of prior investments and of subordinate interim rights to share in enterprise earnings at the discretion of shareholders and elected directors, which allow for retained earnings too.[13]

Accounting regulation does not allow dividends until the earnings actually emerge, and prohibits repaying equity to owners. At the same time, an accounting convention that treats an immediate expenditure as an investment (and thus makes it capitalised as an accounting asset) may completely transform the firm's tax bill (relevant for shareholder value and dividend distribution). Before liquidation, the firm-entity does not repay equity shares at their value, at neither market nor book values. Not only ownership and control, but ownership, control and *management* are separate, and only management can organise the becoming activity of the business and dispose of assets and streams.

Finally, the risk-bearing and taking of business activity relates to the dynamic connection between firm-entity, organisation and management. The whole firm as an entity specialises in risk-bearing, and all the actors are involved in this collective assumption of risk.[14] The decision-making process provided by management, particularly in regard to risk-taking, engages the inner working of a complex organisation in real dynamics, and concerns both the entity and all the undertakers (stakeholders), subject to their different involvements in it.

In fact, managing the firm's risks implies not only the taking and bearing of exogenous risks, but the active management of them.[15] Risks, and the related technological, organisational or strategic alternatives, become endogenous[16] and interact dynamically with the actual economic coordination of the firm-entity as a whole. When uncertain alternatives occur in real dynamics, the primary problem becomes one of deciding what do, and when and how to go about doing it – the problem becomes one of management. In activities such as forecasting consumer wants and their quantity and duration, management does not only bear the risks taken, but also deals with them and thus realises the related incomes. In this active interaction with the emerging economic process of the firm, alternatives could be only vaguely defined and never actually realised. Alternatives constitute simply opportunities for gains or losses. The *management* of the actual *organisation* in the dynamic *process* of becoming transforms them into emerging results. This is why these three components are the inner *constituents* of the firm-entity as a whole.

In framing and analysing the economic and monetary process specific to the firm, incomplete-contracts economics is biased by the entrepreneurial role assigned to the owners (stockholders).

Regarding the purpose of business activity, this role maintains some "principle of maximisation"[17] of the (residual or specific) *income to the owners*, since the firm is nothing other than an owners' device. The distinction between *expected* and *actual* results thus melts away and features such as bounded rationality and uncertainties are greatly reduced.

Concerning the firm as economic organisation, this role implies the design of a rigid internal structure, crucially based on incomplete contracts; opportunism and complete external non-verifiability (*ex ante* and *ex post*) are the sole justifications for the design itself.

The (specific or residual) *income to the firm* is essentially linked to ownership and to the related investments of proprietors, enhancing their property value.[18] This dependence makes the framework unable to link this income either to the firm revenues (i.e., to the interaction between the entity and ultimate consumers), or to other kinds of investment financing, especially longer-term credits. This makes it difficult to understand why taxes are paid on streams other than dividends and capital gains. Finally, since income flow to ownership justifies the firm as an economic process, what about losses? What about constraints on dividends? What about the double taxation of dividends and of net income to the firm?

The ownership connection and the related rigid inner structure weaken the understanding of the interaction between the firm and the "markets" (both product and factor markets), since business incomes and results – factually generated by the firm's dynamic connection between costs and revenues – are seen to emerge within bounding external options that strictly determine them. The special economic *activity* of the firm, therefore, seems to disappear, as well as its *active* role in creating and allocating resources in real dynamics and complexity.

However original and valuable incomplete-contracts economics may be, the dualism between the firm and the market is problematic. The active role of institutions in the special economic process of the firm relates only to the notion of *property rights*, which does not fit the law, economics, and accounting of the firm. Although the merit of this approach is to point to the institutional connection between the firm and the management of its special economic process, the answer provided is a misleading frame of reference and analysis. The same can be said of the notion of *opportunism* that, while pointing to the organisational connection between the firm and the actual organisation, misleadingly overlooks its economic process in real dynamics and complexity. To say the least, an understanding of the economic nature of the firm as institution and organisation has yet to be provided.

## Accounting system, firm dynamics and "the institutional structure of production"

### *The entity view of the firm: the firm as a whole and a dynamic system*

The framework provided by incomplete-contracts economics appears as a genuine but bounded innovation for the economic theory of the firm. It critically stands on strong (and dismal) hypotheses about individuals' motives and behaviour, such as *opportunism* and *hold-up*, which are the sole ultimate justification for the related analysis. A circular reasoning seems to be established between these hypotheses and the critical consequences they imply.

Recent advances question this frame of reference and analysis, seeking new understandings of the economic nature of the firm as institution and organisation. Here incompleteness of contracts implies that *off-contracting* features play a crucial role in coordination, and the scope of contractual analysis tends, therefore, to shrink (Favereau 1997). In particular, Baker, Gibbons and Murphy stress the *relational*, *enduring*, and *informal* features of the special business activity called the firm. Rajan and Zingales are also concerned with protecting the integrity of the firm from the "dark side of ownership." According to them, the whole firm performs an autonomous economic coordination between all the actors involved.

Following these advances, *management* of an ongoing, enduring *activity* emerges as the primary way to overcome the alleged dualism between the firm and the market.[19] In both his classic and his recent articles, Coase insists on the firm as economic coordination, but disregards *opportunism* and *hold-up* as proper foundations for the economic theory of the firm (Coase 1988). According to Coase, management direction appears to coordinate *by fiat* not only employees, but all the resources involved, including the equity finances. He finally links the special economic nature of the firm to: (a) the *economic* distinction between firm and market, since the firm supersedes the price system; (b) the *active* role of legal institutions in this special economics of the firm; (c) the inner working of organisation and especially of the accounting system as distinctive firm's features for superseding of the price system (Coase 1990).

Notwithstanding that recent developments renew and further develop the transactional and institutional perspective, their insights are still constrained by modelling the firm with reference to an odd *nexus of prices, contracts or property rights*. This contractarian view limits the real revolutionary force of ideas such as bounded rationality and of the critique of the equilibrium firm Simon (1991; 1997) developed starting from the firm as an economic system, and quoting Commons and old institutionalism as a key thought-provoking precedent to his own theorising. The recent advances still frame the economic analysis of the single firm with equilibrium, as incomplete-contracts economics does. Equilibrium designs at best an *ex ante* mode of functioning based on atomistic elements or individuals, and neglects all the implications of real dynamics and complexity for the whole firm.[20] Finally, framing the single firm with equilibrium weakens such an understanding of its economic nature as institution and organisation that real dynamics and complexity would enhance.

On the contrary, Simon (1991; 1997) stresses the need for a new theory grounded on *active* firms instead of paramount efficient markets. Starting from this emerging notion of the firm as an *entity*, Simon discusses the production process and the related web of products' selling transactions. His emphasis is there on the notion of a *dynamic system*, with its key feature of *feed-back* effects, far away from the equilibrium mechanism advocated by neoclassical economics:

> The common and understandable practice of pricing by *marking up* costs assures liquidity, at least in the short run, if only there is at least a modest

> base of fixed costs.[21] The *adjustment of production rates* to sales holds price margins within a moderate range without excessive absorption of cash by inventories. *All of this has little or nothing to do with the usual theorems of optimal pricing and production rates*. A simple *feedback* of price, inventory and sales information adjusts production and prices and maintains a *tolerable steady state* over considerable intervals of time without any close calculation of margins or optima.
>
> (Simon 1997: 37, italics added).

From this perspective, Simon shakes both the profit maximisation and the equilibrium framework, both essentially static, as foundations of economic theory and analysis of the firm. Instead, he rediscovers a steady view of the economy flowing, nevertheless enhanced with a decisive dynamic glamour. The firm as an entity is hence understood as a *dynamic system of interactions, interdependencies, complementarities*, not only contractual or bargaining, located in time and space, and different in nature from any equilibrium nexus of prices, contracts or property rights.

From this perspective, the firm may be understood as an *entity*, as a special *field* providing an enduring pattern of interactions between all the parts involved in the actual process of becoming for its business activity. The firm as economic entity exists and functions as a *whole*, as a dynamic system of interacting parts, coordinated by human intelligence and effort (management), committed to, and involved in, an economic undertaking (*managed economic system*). By parts we mean all resources, whether material or personal, financial or economic, tangible or intangible, provided by all the actors or individuals that undertake the business (undertakers or stakeholders). The special, economic activity of the firm-entity is usually carried on through the medium of *transactions*, which are relational exchanges of goods and/or services for a consideration, and of *combinations* of the parts in going processes and working structures. By an economic undertaking, we mean the implication of parts, coordinated by management, towards the performance of a special, becoming activity that is organised in terms of fulfilling a collective purpose.

In this context, the firm-entity *exists* as a whole in real dynamics and complexity, because of *modes of functioning* that jointly *constitute* its special process of becoming. These modes are called "constituents" and allow the whole to be different than the mere additive resultant of the parts. Any of the single parts can be, and in time will be, changed or replaced, and yet the firm as a whole will go on without interruption and without abrupt change in functional characteristics. Reciprocally, people undertake firms business activity as providers of resources. They know and play their roles without losing their own autonomy (and liberty).[22] As such, the constituted whole *functions* apart enough from the parts who compose it (actually, in a greater order), but it *does not perform* and *fulfil* apart from them (this point expands upon Stauss 1944, reprinted in this volume).[23]

The firm is hence one entity *functioning* through some constituents and *performing* through the medium of various interacting parts having heterogeneous

responsibilities to and involvement in it. The firm-entity combines the parts in such a manner as to give the parts purpose and meaning, and thereby it influences the parts and makes them components of a constituted whole. It involves and commits actors or individuals both as providers of resources and as undertakers of becoming economic activity seeking for intended results.[24] The whole, then, influences their behaviour, prompts, frames and enhances it at the same time, but does not determine it.

From the transactional and institutional viewpoint, as an undertaking of any size or complexity, the entity is not merely an agency for the stockholders alone. Its institutional structure of production is more than a bargaining nexus of (ill-defined) property rights. On the contrary, no one "owns" a business entity. Instead of owners, there are various providers of resources, there are various actors or individuals which are involved in different types of *transactions with* and *combinations into* the entity.[25] The firm-entity is then active and productive – but is so because it is managed and organised, not because it is "owned." Its value depends, not on wealth of resources passively held, but on their managed system dealing with the *flow* of relationships generating incomes, that is, on dynamics and process. Stop the dynamics, and its value as an entity disappears.

### *The accounting understanding of the nature of the firm as an entity*

In understanding this entity system, the accounting system may provide the next theoretical step.[26] In a dynamic context, accounting can provide an original view of the special economics of the firm. It may further constitute the way allowing the firm to supersede the price system and to be different from a nexus of contracts or prices, since:

- accounting deals with the firm as an entity: business activity is seen as a dynamic concern and the accounting system has to report on it.
- This reporting is especially concerned with the representation of business incomes to the firm (so-called gross and net earnings). It is for this income that the entity is accounted for (accountability).
- The accounting system, therefore, becomes a mode of knowing, organising and regulating the economic and monetary process belonging to the firm as an entity.

In this context, as suggested by the early accounting and economics of the firm,[27] the accounting system may be understood as one constituent, one mode of functioning for the firm-entity as a whole. Accounting may provide an alternative frame of analysis allowing us to grasp the special economics of the whole firm and to overcome the alleged dualism between the firm and the market. Accounting, in fact, stands on general principles that we can summarise in the following:

(a) The entity principle: the business firm is an entity and a going concern, autonomous from whichever stakeholders (including shareholders).

(b) The matching principle: a special method to link the economic and monetary entity's streams to the reference period.
(c) The historical or invested cost principle: a special method to recognise and estimate actual business activities as assets and liabilities, costs and revenues.

These principles were and still are in question, but they are, at present, the principles generally accepted to represent the firm in accounting. They constitute the accounting view we are looking for.[28]

These principles are often questioned because of the special view of the firm they imply. As recognised by leading accounting theorists (Zappa, Schmalenbach, Nicklisch, Littleton, Ijiri, Anthony), the accounting view deals with the firm as an entity and with its events, resources, and transactions in real dynamics and complexity. According to Shubik (1993), time and uncertainties have essentially disappeared from the apotheosis of *price system* driven by equilibrium framework, but they remain the concerns of everyday business activity. The problems of how to account for their influence in the ongoing economic *process* are central to the development of accounting and lead to the original accounting view of the special economics of the firm seen as an entity.

The entity's real dynamics imply uncertainties, imperfect knowledge, potential and actual mistakes and misorganisation. Dynamics inscribes business activity into a special economic *process* of becoming, and accounting has to cope with this entity's process. The accounting view then represents the economic and monetary process of the whole firm in a very different way the equilibrium framework does, as we see in Table 14.1.

However instantaneous or stationary, an equilibrium framework implies synchronicity of costs and revenues, conceived as simultaneous cash flows imputed to each product separately. Every transaction is indeed closed by a monetary flow that is a market price.[29] A limiting identity is usually assumed between contractual transaction, market exchange and market price, allowing the independence of periods. No economic system other than the price system exists.

On the contrary, the accounting system is grounded on events, resources, transactions and combinations as they actually happened, engaged or committed by a given firm. Such a system deals with the whole firm as an entity, melts real and monetary matters, and makes periods interdependent. Starting from actual monetary transactions between one entity and its world,[30] the

*Table 14.1* Equilibrium and accounting frameworks for the economic and monetary process of the firm

| | *Assets* | *Liabilities* |
|---|---|---|
| *Equilibrium frame of analysis* | future monetary *entries* discounted | claims on future monetary *exits* |
| *Accounting frame of analysis* | actual monetary *exits* (expenditure) capitalised | advances on future monetary *entries* (concerned with real dynamics) |

accounting system fills in the inner and inter-temporal allocations of prices-related and income-related *accounting values*.

Hence the accounting system provides a representation of the economic and monetary process of the firm starting from the dynamic connection between two autonomous and interdependent patterns of economic streams. These relate either to the costs' side or to the revenues' side of the becoming business activity of investing, producing and selling.

This accounting representation focuses on the resources' application for which the entity is accounted for (*accountability*). The dynamic connection between costs and revenues constitutes a more reliable and verifiable aggregation allowing the estimation of the actual performance generated by the firm-entity as a whole during the period of reference.

As a result, the accounting notions of costs, assets and selling prices do not agree with the notions provided by the equilibrium framework:

- Accounting costs *match* either the actual products sold, or the reference period (costs of being in business) for income representation.[31]
- In most cases, accounting values for capitalised costs (i.e. assets) do not discount future entries (i.e. results) from related business activity, since the actual determination of these emerging results (if any) is the primary purpose for which that entity is accounted for.
- Selling prices do not constitute merely informative signals to actors engaged in equilibrium adjustments, but payment inflows necessary in order to recover incurred costs, repay debts for investment and working capital and meet other obligations and claims in real dynamics.[32]

In this context, most accounting assets relate to *specific* financial *exits* (expenditures) that accounting rules capitalise. These assets make the whole firm path-dependent in a potentially irreversible way, since expenditures might never be recovered by future selling revenues.[33] At the same time, contrary to the myth of the lonely entrepreneur, shareholder equity does not appear here as a claim on the assets' *values*, but as an advancing of invested *costs*. It is the *source of funds* that is waiting for eventual dividends. But accounting *earnings*, as ultimate outcomes of the working of the accounting system, provide an upper legal constraint on those dividends. At the same time, the accounting system reveals the tax base derived from the income generated by the business entity. Accounting participates therefore in the active role of legal institutions in business income allocation.

To sum up, at least three features distinguish the accounting system from the price system:[34]

1 The accounting system is based on accruals: it grounds on financial transfers and not only on cash. A "veil" is therefore provided over every monetary price and the inner economic process, since, for instance, relevant transactions or events are recognised no matter if they are paid in cash or not;

2 Accounting's special logic of regulating, organising and representing the firm's activity (related to overheads, accruals and other accounting values) further implies the *active* role of the whole firm in real dynamics and complexity;
3 As a consequence, the accounting system provides relevant figures for both dividend and other legally-enforced constraints, as well as for paying taxes on the business income generated by the whole firm.

Real dynamics and complexity, as well as the separation between ownership, management and control (Berle, Littleton) imply the entity view on the firm. This view recognises the old-fashioned proprietary view as irrevocably lost, all the more so in regard of complex legal-economic webs that currently constitute the institutional bundling for the business entities. *Even to protect shareholders*, we need a new kind of control, different from the irrevocably lost ownership sovereignty.[35]

Given real dynamics and the separation between ownership, management and control, the accounting representation of the firm does not put together rents, properties, and related claims on (well defined) property rights, but deals with actual revenues, invested costs and funds advanced for them. The accounting system cannot and will not establish *expected* results for certain undertakers, but the actual entity's performance represented in a reliable and traceable way, as shown in Table 14.2.

In summary, by means of its dynamic entity view, the accounting system defines the *business incomes to the firm as an entity* (Zappa) and plays an active role on their allocation, allowing the special economics of the firm to supersede the price system.

According to Baker, Gibbons and Murphy, the business income stemming from the entity's coordination essentially relates to the entity's ability to continue performing.[36] As long as actual coordination goes on, it involves a relational and informal component beyond contract for each resource provider. This involving commitment implies, for example, bearing some business risks involved in the business activity, i.e., the risks taken by management and implemented by the actual organisation in real dynamics. The commitment modifies income allocation, attaching a relational, specific income to every provider, subject to different arrangements. Furthermore, the institutional structure shaping the business activity requires *prior* recovery of invested costs for the whole entity, costs that ought to include fair equity interest for providers of equity finances (holders of equity shares), based on actual funds provided in the past.

*Table 14.2* Proprietary and entity views of the firm

| | *Incomes* | *Assets* | *Liabilities* |
|---|---|---|---|
| *Proprietary view* | Rents or quasi-rents | Properties | Claims on properties |
| *Entity view* | Business incomes to the firm (based on actual revenues and costs) | Costs invested | Revenues advanced (concerned with real dynamics) |

Thus, rather than the so-called "ownership structure" of the firm, it is the enduring existence and financial viability of the whole firm that is fundamental. The accounting system constitutes one of the modes of functioning, i.e., it is a constituent, of the firm-entity. It relates jointly to the basic learning, diffusion, and information processes (the epistemic structure of production), to the closeness of products, technologies, resources and internal organisation (organisational structure), and to the role of working rules and norms dealing also with financial matters and regulation (institutional structure).

The accounting system, therefore, with limitations, allows us to *know, organise and regulate* the special process of becoming as it emerges into the firm as a whole. This new understanding of the accounting role calls for a further analysis of such structures that characterise the economic nature of the firm as an entity.

## The firm as an entity

### *The firm-entity as an economic organisation and institution*

How can the nature of the firm as economic organisation and institution be understood?

The alleged dualism between the firm and the market constitutes a very limited representation of the dynamic connection between management, organisation, and the accounting system. Each of these three *inner* constituents of the firm-entity is quite distinct in nature from the others. Management, as the capacity to induce and maintain actions, interactions and activities, complements the possible modes of existence of the organisation, creating structures and establishing order and coordination. This interaction between management and organisation exists and evolves in the *actual* process of becoming. This process may be, with limitations, represented, organised and shaped by the accounting system.[37]

In this context, the firm-entity exists as the *actual economic coordination* oriented by management by means of the organisation that the accounting system, with limitations, represents. Management blends with the organisation to actualise the special process of becoming of the business activity. In such a way, the whole process of becoming of the firm as one entity is not confined to some disparate series of events, resources and transactions. The firm is not simply a legal fiction standing for the nexus of contracts between separate individuals or proprietors. There is no activity such as a business undertaking unless management, organisation and ongoing process are present. They *constitute* the inner cohesion for the whole firm as an entity, i.e. the managed economic system that allows the business activity to overcome hazard and uncertainties under real dynamics and complexity.

This actual economic coordination that characterises the nature of the firm is neither blind nor spontaneous, but allows conscious and creative actions. The entity's economic *activity* is indeed purposive in real dynamics and complexity. It is a *becoming* concern, since orienting this activity concerns at least every

resources provider, including employees, managers and final consumers, as long as their resources are committed to the firm-entity and engaged in the inner working of its organisation. In one sentence, as long as the entity's economic coordination *goes on*, the entity's economic activity *becomes*.

Contrary to incomplete-contracts economics, most interdependencies and complementarities do not relate here to the *contractual* interaction between two (or more) actors (*nexus of contracts*), but to the involvement of each undertaker in the special economic process of becoming as it emerges into the whole firm. This involvement may relate either to the revenues' side or to the costs' side of the becoming economic activity of the firm-entity, according to specific organisational, institutional and epistemic conditions.

Undertakers then provide the firm with a purposive framework that modifies each constituent in the *actual economic coordination*, and transforms it into a *becoming economic activity* seeking satisfactory realisations. This is why both incomes and results and the undertakers are the outside constituents of the firm as an entity.

From this perspective, the economic notion of firm as an entity, its being a whole, refers at the same time to the *inner* relationship (point 1 below) between management, organisation and the going process (defining the entity as actual economic coordination), and to the *outer* relationship (point 2 below) between the functional entity and its undertakers (defining it as a becoming economic activity), concerned with the creation and allocation of emerging business incomes and results. These fundamental relationships define the firm as an economic organisation and institution:

1 The inner relationship of the entity as actual economic coordination, where the management direction coordinates and focuses the inner working of organisation into the going economic and monetary process. This *managed system* defines the entity's economic nature, and generates the becoming economic activity of the whole firm.
2 This activity is outlined by the outer further relationship, which considers the business entity from an external viewpoint. Emphasis is here on the dynamic connection between the functional entity seeking for realisations of its activity (*intended* and *actual* incomes and results), and the undertakers looking for them. Management coordinates this activity with the fiduciary authority conferred by all the undertakers and according to institutional arrangements.

Even though undertakers have purposive *expectations* of an entity's results, whether promised or residual (outer relationship – 2), *actual* fulfilment of them depends on the undertakers' involvement in the inner working of the organisation coordinated by management given its delegated authority (inner relationship – 1).[38]

From the transactional and institutional perspective, notions such as economic organisation (organisational structure of production), or institution (institutional structure of production), or knowledge (epistemic structure of production), relate to the fundamental *interactions* between *constituents* (functional

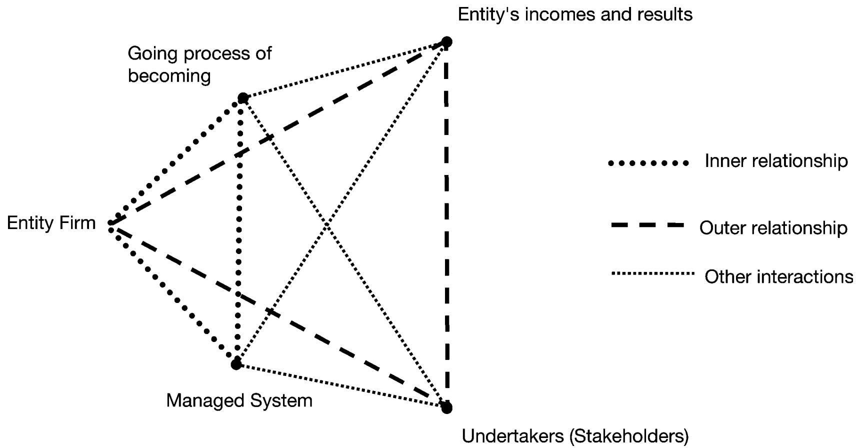


*Figure 14.2* The firm-entity as an institution and an economic organisation.

modes of existence) as we can experience and understand them in real dynamics and complexity.

In particular, *property rights* are at best an incomplete and at worst a biased acknowledgement of the active impact of legal institutions on the whole entity and on its management and going process of becoming (institutional interaction). Similarly, *opportunism* can only be an incomplete and surely dismal acknowledgement of the active role of the inner working of the organisation in the whole entity and in its special process and emerging results (organisational interaction). Finally, *bounded rationality* is a valuable but underdeveloped idea for grasping the ways in which actors engage in actual business decisions and activities under real dynamics and complexity (epistemic interaction).

Arrow (1986) argues with Coase about the idea that *property rights* alone can provide social welfare under bounded rationality, and replies with the *price system* seen as a frame of competitive interactions. In fact, under incomplete and imperfect markets, only firms as *active* entities provide sustainable conditions for effective business activity of investing, producing and selling. Put differently, the "contractarian firm" does not appear to have emerged from the making of reality, but from the day-dreaming field of the price system.[39] Instead, under real dynamics and complexity, the leading characters on the socio-economic scene are firms as entities.[40] Neither *property rights* nor *markets* (i.e., the price system under competitive conditions) taken alone ensure effective economic production. Just as Simon's creature floating down to Earth from Mars does, the entity approach understands firms as connected by a changing network of communications and transactions that we know as markets, and "surely [the firms] would appear to be the active elements in the scene" (Simon 1997: 35).[41]

In the light of this predominance (and organisations), markets may be considered as threading environments that link firms-entities to each other and to undertakers. Markets may (or may not) determine this *outer* interaction, but

never the emerging process of becoming for the entity's activity. The firm-entity is not a *price* system (i.e. the reacting resultant of competitive conditions), but a *managed* system. It is the *actual economic coordination* that characterises the whole firm as an entity, defining its special economic nature. This view of the whole economic system relegates the notion of market to explicitly designed arrangements, like the "stock market" and so on. Most of the so-called "market exchanges" are instead outer transactions and dynamic interactions between the entity and its world.

According to this transactional and institutional viewpoint, (properly defined) property rights alone do not *constitute* the institutions shaping the economic activity of the firm, since so do all norms, rules and conventions designed for and actually applied in business activities, including the accounting-based standard representation for income taxes and dividends' constraint. All these working regulations play an active role as institutional structure of production in the special economics of the firm, and constitute its legal-economic framework (or system).[42]

Legal institutions are especially concerned with the outer interactions between management and undertakers, and between undertakers and the entity as a whole. They frame the business activity, shaping and modifying the inner relationship between the entity and the management of its going process. In this context, management direction coordinates *by fiat* not only employees (labour interactions), but all the resources involved, including the equity finances. Management, however shaped and modified, never plays the capitalistic hero cherished by incomplete-contracts economics. The firm-entity is not a bundle of property rights, however these may be defined. In real dynamics, the inner working of the organisation and the actual going process of the firm as a whole constitute the performing but limiting factors for the managerial active role.

### *The role played by the accounting system in business governance and disclosure*

In this dynamic context, the accounting system *functions* as:

- a mode of knowing (and representing) the going process and its *actual*, emerging results (epistemic role);
- a mode of regulating the entity's activity and shaping its special process of becoming for the undertakers and the law, including fiscal policy, especially concerned with the structured allocation of *intended* results (institutional role);
- a mode of organising the becoming economic activity in accordance with the going economic and monetary process of the whole firm (organisational role).

The accounting system constitutes one of the modes of functioning (constituents) of the becoming activity of the firm-entity, as *cognitive tool* and *organisational instrument* for management, and as *working norm* (and *rule*) related to the institutional structure of production.[43] The accounting system deals with the effect on

each entity of events, transactions and combinations which take place in its special process of becoming. In real dynamics and complexity, however, the accounting system does not determine the process that it has to represent, organise and regulate. Instead, the accounting system and the going process are to be distinguished. For example, the accounting system – as one of the entity's modes of functioning – involves the use of money as a symbol, whilst transactions between one entity and its world involve the use of money as a medium of exchanges.[44]

Since the accounting system does not determine the functioning of the related economic and monetary process, its *epistemic* role does not fit with any universal "principle of value maximisation." The business firm may seek incomes, but this quest does not explain the fundamentals of its special working. According to Anthony (1960; 1983), the accounting framework relies instead on concepts and rules making the *history* and *life* of the firm-entity understandable *synthetically* under bounded rationality. Accounting is concerned here with the capacity of the firm-entity to go on, to become and to *satisficingly* fulfil in a situated, changing context.[45]

This capacity relates to the *recovery of invested costs* in real dynamics, as well as to the satisficing creation and allocation of business incomes for all the undertakers, including a satisficing *equity interest* for shareholders as providers of equity finances.[46] Calculated on the basis of *funds actually committed* in the past, and compensated with actual dividends paid out, this interest recognises shareholders' equity as a special dynamic source of funds and may constitute the accounting way to bridge intended and actual results in real dynamics. This comes to the following operational definition of the net income to the firm, i.e., the net income after shareholders' equity interest:[47]

- If this kind of "net income" is positive, it can indicate *over-profits* (or *quasi-rents*, in the terms of the new theories of the firm), that, in the pattern of entity theory, can be distributed not only to shareholders, but to stakeholders and even the community (corporate gifts);
- If this "net income" is negative, it becomes an inefficiency signal, with an implicit menace in terms of survival, since financial stockholders may not accept this state of affairs. This is not, however, a signal of failure or "technical" financial distress.

This calculation may complete the accounting role as a reasonably aggregated, *reliable* mode of knowledge and governance. It keeps the accounting system logically independent from the price system, and therefore makes it suitable for inside settling of conflicting interests, for guiding the proper formation of prices on the stock market in time, and for addressing and evaluating both the entity's and its managerial performances.

In sum, the accounting system constitutes a mode of knowing, organising and shaping the entity's special economic and monetary process that emerges and evolves into the business entity as a whole.

## Conclusion

Following the genuine insight of Coase, Simon's Martian may delve into accounting in order to understand the special economics of the firm. Although incomplete-contracts economics is lacking in the Martian approach, arguably because of the alleged dualism between the firm and the market, a new transactional and institutional perspective can deal with the firm. This essay outlines this development starting from Coase, Shubik and Simon, but also from the recent advances provided by Baker, Gibbons and Murphy and Rajan and Zingales.

The new perspective recognises the firm as an entity, as a special field providing an enduring pattern of interactions between all the parts involved into the actual process of becoming for the whole firm. The firm as economic entity exists and functions as a *whole*, as a dynamic system of interacting parts, coordinated by human intelligence and effort (management), committed to, and involved in, an economic undertaking.

Therefore, the interaction of the parts is not sufficient to understand the durable existence and functioning of the firm-entity. In real dynamics, the firm-entity further appears as a special business *activity*, analysable in terms of modes of functioning (constituents) which jointly constitute its actual process of becoming as a whole. Each mode is analysable within the entity as an autonomous, active component, and "in its *interaction* with" ("playing an *active role* in") the *economic and monetary process* with which this special activity is concerned. In this transactional and institutional framework, the alleged dualism between the firm and the market disappears, and the firm-entity as a whole can *exist* and *function* apart enough from the parts that compose it, but does not *perform* and *fulfil* apart from them.

In this context, the accounting system is seen as one of these modes of functioning of the whole firm as an entity. It relates to the epistemic, organisational and institutional *structures* (or systems) of production, i.e. it constitutes a mode of knowing, organising and shaping the actual, going process of the firm-entity, dealing with real dynamics and complexity.

In particular, it provides a *reliable* representation of actual transactions and combinations, events and emerging results. Thus, it may trace and estimate the becoming economic activity in real dynamics, and play an active role in the entity's economic and monetary process. The accounting system, indeed, will allow the special economics of the firm to *supersede* the price system into its special process of creating and allocating resources.

Whilst incomplete-contracts economics allows firms to emerge in the theoretical framework provided by the price system, this transactional and institutional perspective seeks in the making of reality to discover the role of economic entities in economy and society. So, even though the firm-entity does not act as individuals do, it becomes an integral part of society as a whole, having its own autonomy, enduring existence, and function as an *economic organisation and institution*. As such it is responsible towards the totality of human experience and of the natural world. Its powerful impact on individual liberty and life ultimately asks for

a "conscious island" of concerted, *managing* authority, looking forward to serve not only its undertakers, but the whole human community.

## Appendix: couples of contrasting notions

| *Incomplete contracts economics* | *Transactional and institutional perspective* |
|---|---|
| Equilibrium | System |
| Dualism between firm and market | Five-terms system |
| Firm as a nexus of prices, contracts of property rights | Firm as an entity |
| Opportunism and hold-up | Structures, epistemic, organizational of institutional |
| Property rights | Working rules and norms |
| Opportunism | Inner working of organization |
| Game theory-based bargaining | Accounting system |
| Separation between property and control | Property, management, and control |
| Power bounded by outside options | Authority based on fiduciary responsibilities |
| Ownership sovereignty | Management coordination |
| Some principle of maximization or economizing | Satisficing based on bounded rationality |
| Net revenues as rents or quasi-rents | Business incomes to the firm |
| Accounting proprietary view | Accounting entity view |

## Glossary

**Constituents** Selected *functional modes of existence* of the system, which jointly constitute the dynamic process of becoming of the whole. They allow the whole to be different from the mere additive resultant of the parts. As such the constituted whole *exists* and *functions* apart enough from the parts who compose it (actually, on a greater order), but it *does not perform* and *fulfil* apart from them.

**Dynamics (and Complexity)** The actual, conditioned, becoming *relationship* as it goes on and evolves in real time and complexity (implying uncertainties, imperfect knowledge, potential and actual mistakes and misorganization). Dynamics is hazardous. Real dynamics stands for dynamics of the reality in its making. The firm's dynamics entails the special process of becoming of the firm as a constituted whole.

**Firm-entity** The firm as an enduring pattern, or whole, experienced and understood as a system with special dynamics and complexity. The firm is hence an entity *functioning* through some constituents and *performing* through the medium of various parts having heterogeneous responsibilities in and interactions with it.

**Interaction** Logical and functional nexus between two (or more) parts (or constituents) of the system as a whole.

**Mode of existence**, Mode of functioning see *Constituents*, *Firm-entity*

**Playing a role in** see *Interaction*

**Parts** The simplest elements of a whole, for example the resources (provided by actors or individuals) committed to the firm-entity. Their identity could not be fundamental for the whole as it *exists*, even though it could become fundamental into its actual process of *becoming*. The whole influences the behaviour of the parts, but does not determine it.

**Process of becoming** The relatively orderly and recognizable flow of *activity* that binds the present, the past, and the future in dynamic evolution. Whenever it presents itself to our experience and understanding, we find in it wholes, relationships and systems. See also *Dynamics*, *Relationship*

**Relationship** Logical and dynamical interaction between *three* constituents of the system as a whole.

**System** A frame of autonomous and yet mutually necessary terms, jointly defining a whole. Terms could be both parts and constituents.

**Undertakers** In a wide meaning, all beings involved in or concerned with the *activity* of the firm. In regard to analysis, all actors or individuals having heterogeneous responsibilities in and interactions with this activity in some relevant way. They participate in but do not belong to the firm.

**Whole** see *Firm-entity*, *system*

## Acknowledgment

This chapter is based on Biondi (2005), and partially draws on Biondi (2006) – by Routledge.

## Notes

1 According to the "as if" epistemological defense, also known as methodological irrealism, provided by Friedman (1953). cf. also Machlup (1946: 534–5) or (1967: 6–7). Instead of predictions of actual behaviour, the utmost purpose for a theory may be to provide framework and perspectives to understand and represent phenomena as we experience them.

2 Baker, Gibbons and Murphy (2001) also stress the difference by nature between the firm and the market.

3 As recent literature will be quoted later in the article, references could be made here both to old institutional economics in Europe and USA, and to some theories of business economics and accounting especially developed in USA, Germany and Italy. About the dynamic interaction between a whole, and its parts and constituents, Copeland (1927) quotes A.N. Whitehead, Lloyd Morgan, and H. Bergson. In this volume, cf. also Canziani, Kirat and Gindis, providing further references.

4 On the contrary, each firm ought to appear as a special case, marked out by its own peculiarities (including actual resultants of behaviour), but otherwise exhibiting the general modes of functioning that jointly constitute each firm.

5 See below, the third section.

6 This is the title of Coase's Nobel Conference.

7 An analogous argument could be developed about the dualism between the firm and the proprietors (shareholders), as well as between the dedicated agency and its principals.

8 Ijiri (1967: 58–64) and (1975: 183–6) provides a perceptive analysis of the different *imputation* logics underlying accounting and economics. He explains how simple production processes melt the straightforward logical chain of cost, quantity, and selling price for each product separately required by the *black box* firm. In such cases, the black box requires some further conditions (epistemic or organizational) external to the accounting system and provided by the price system. The dynamic accounting approach, instead, based on historical or invested costs, does not require the *separability, stability and uniqueness* of that logical chain (cf. also Biondi 2003: 19–21). Its approach is more aggregating than individualistic, its figures are more actual amounts recognized than market or discounted values, it prefers traceable and reliable discretionary methods (accruals) for optimising or market-to-models methods.
Kirman (1997) enhances the economic viewpoint by dealing with the evolving network of agents involved. His conclusions are thus very different from usual approaches grounded on separate or representative actors. His analysis treats in particular the monetary profit, actual and *cumulated*, generated by each transaction separately.

9 The notion of constituent is inspired by Whitehead (1932: 131); cf. as well as Royce's (1914) order-system of relations as modes of action.

10 The Incorporation Act also establishes this ability from the legal viewpoint, see Gindis and Manfrin (both in this volume). For historical legal perspectives on the firm as an entity, see Avy-Yonah & Sivan, and Kirat (both in this volume). The chief reference for law and institutional economics of the firm is still the work of A.A. Berle, twice reprinted in this volume.

11 From an old-institutional viewpoint, TCT provides an approach based on transactions *without* going concerns. Further distinctions could be made between the classic Williamson (1975) and later developments within incomplete-contracts economics.

12 In accounting, this figure influences the old-fashioned proprietary theory of the firm, cf. Gynther (1967) and Sprouse (1957).

13 A further obligation of owners, the role of guarantors for equity provided against losses beyond a legal limit, is currently avoided by business practices.

14 Both as *passive* assets' aggregation and as *active* impact of its existence as a whole.

15 Cf. also Littleton (1928).

16 On growth real options as endogenous, cf. Zingales (2000).

17 Or "economising," in Williamson's term.

18 An individualistic property rights approach requires that each proprietor should hold his asset(s) alone in order for the exit threat to be real.

19 cf. especially Baker, Gibbons and Murphy (2002: 73–4); Crémer (1986).

20 In fact, Pareto (1906) deals with the economic *system*, but there is only a system in equilibrium, having emergent but spontaneous order. Nevertheless, his general equilibrium framework allows for the intended purpose of firms (profit) to be not actually realized, and for firms to be different, subject to general competitive conditions. Furthermore, Pareto ultimately disregards prices, for he relegates them to accessory tools, to conceive a full and sole system of preferences and constraints, where individual values (preferences) cannot be reduced to prices. This framework might therefore be understood as a partial and incomplete framing for the whole economic system. Even so, the equilibrium framework for the single firm is disregarded here.

21 Simon applies here the following assumptions: (a) fixed costs are completely sunk, i.e. already paid-out as actual past expenditures capitalised into accounting assets; (b) production is accomplished under the maximal level of utilisation for that bundle of fixed (invested) costs.

22 See also Ruben (1983)'s distinction between being a part or a member of social wholes, discussed by Gindis (this volume).

23 Its "frontiers," then, are established by the *inner* relationship concerning its *actual economic coordination*, constituted by management, organisation and its special process, as contrasted to the *outer* constituents that are the undertakers (stakeholders), and its incomes and results.

24 In this context, the ultimate consumers are undertakers as well, seeking final products for a consideration (usually the money transfer called price closing the transaction).

25 This smooth distinction between transactions and combinations depends on the difference between entity and markets. In such a context, financing and work activities are *combined* into the entity, and the interactions between the entity and the providers of these resources are quite different in nature from *market* exchanges. Schumpeter (1912) used the word "combination" (*Kombination*) to describe the new productive process introduced by the innovative entrepreneur by means of his firm. K. Polanyi suggested we should overcome "our obsolete market mentality" and realise that money (related to capital and credit), land, and labour are "fictitious commodities."

26 At this level of generalisation, the notion of "accounting system" stands for (i) operations' costs and managerial accounting related to the organisational structure; (ii) financial accounting and tax accounting related to the institutional structure; (iii) disclosed accounting information (related to managerial and financial accounting and integrated into the epistemic structure).

27 Cf. especially G. Zappa and A.C. Littleton. Canziani (this volume) speaks about the continental tradition of business economics and accounting promoted by Schmalenbach, Nicklisch and Zappa. Berle (1947, reprinted in this volume) recognised the importance of accounting for his theory of enterprise entity.

28 Accounting notions stand here for the usual notions and rules framed by the General Accepted Accounting Principles (GAAP) and by the historical cost approach. In this context, *accruals* generate an original view and frame of analysis for the economic process of the firm-entity as a whole.
Littleton (1953: 24) adds a fourth principle of enterprise service: "Business enterprises are accepted and used because they perform effective economic function in supplying goods (for living) and employment (for earning)."
Even though recent developments enhance the role of discounting-based values for some special cases, the genuine originality of the accounting view still remains. In fact, under real dynamics and the separation between ownership, management and control, the accounting system frames and modifies the economic and monetary process of the firm *whichever accounting approach is retained*.

29 "Money talks" in Williamson's (1991) terms.

30 As a method of recognition and measurement, accounting is not grounded on cash receipts and disbursements (cash-basis accounting). In short, either cash transfers or occurrences of credits/debts suffice to record transactions into accounts. For this and other matters, accounting method makes periods inter-dependent.

31 Cf. Anthony (1983, chap. 5: especially p. 124 ff.).

32 Furthermore, selling prices relate to the interaction between the firm-entity and the final consumers, including discriminatory and multiple-prices policies.

33 Thus, *income to the firm-entity* does not relate primarily to stockholder subordinated claims, but primarily to the selling revenues gained from consumers. It relates to satisficing actual results for all the undertakers involved in the business activity.

34 Even though the *fair value* approach includes special values based on discounting and marked-to-models, these values cannot liken the accounting system to the price system. See note 28.

35 In this context, the entity view on accounting regulation is not so much about single assets' valuation as about recognising the firm's actual economic system (and the involved risks and implications) over and above the very thin corporate frontiers. The so-called *off balance sheet* items are not *off* the flow of relationships constituting the firm as an entity.

36 At the same time, the notion of *organisational capital* as suggested by Rajan and Zingales may be reformulated in the accounting frame of analysis, subject to the distinction between intended results (the old-fashioned financial goodwill based on discounting), and the actual, emerging incomes represented by the accounting system.
37 Searching for business income finalises the interaction between management and organisation. The purpose of business income for the firm-entity is partially institutionalised by accounting rules determining earnings, and differs from optimal profit framed by single firm equilibrium.
38 This approximation fails to further link expectations and realisations. While it refines the economic notion of the firm as an entity, it does not deal directly with the field and the ways in which its business activity becomes or its going process actualises. This question should be deepened in a further approximation, related to the entity's life in real dynamics of transactions and combinations, events, including both *accounting system* and *intended and emerging results* as further terms, cf. Biondi (2005: 35ff.) for further details.
39 In spite of the methodological realism Coase (1988: 52–4) advocates.
40 We are neglecting the mystery of other "black boxes" such as the state, families, non-business entities, and of the monetary and financial system.
41 Cf. also Simon (1991: 27, reprinted in this volume).
42 This idea is further developed on this basis by Manfrin (this volume).
43 Following Anthony (1983: 208, note 6), the usual barrier between financial and management accounting is left out and further developed as inner and outer accountability.
44 Cf. Zappa (1937), but also Littleton (1961, *Symbols of Reality*: 226–7); Raby (1959).
45 Anthony directly refers to Simon's approach. Cf. also Arrow (1986).
46 This was argued long ago by Schumpeter (1912: 175–200). A sort of *equity interest* is also calculated by the *Economic Value Added* (EVA) approach to residual income, even though it usually starts from invested capital (capitalised as accounting assets). To be sure, in order to make accounting and discounting-based measures compatible, the simple interest method ought to be preferred.
47 Professor Anthony declares this definition equivalent to his own in our epistolary exchange. In fact, this connection between net income and corporate gift appears to be neglected before Biondi (2002, chap. XV: 341 ff.).

## References


Anthony, Robert N. (1960), "The Trouble With Profit Maximization," *Harvard Business Review*, 38 (6), November–December, reprinted in this volume.

—— (1983), *Tell it like it was: aconceptual framework for financial accounting*, USA: Richard D. Irwin Inc.

Arrow, Kenneth J. (1986), "Economic theory and the hypothesis of rationality," *Journal of Business*, 59 (4); pt. 2 in J. Eatwell, M. Milgate and P. Newman (eds), *The New Palgrave. A Dictionary of Economics*, London: Macmillan, 1987: 69–75.

Baker, George, Gibbons, Robert and Murphy, Kevin J. (2001), "Bringing the market inside the firm?," *American Economic Review*, 91 (2), Papers&Proceedings, May: 212–18.

—— (2002), "Relational contracts and the theory of the firm," *Quarterly Journal of Economics*, 117 (1), February: 39–84.

Berle, Adolf A. Jr. and Means, Gardiner C. (1932), *The modern corporation and private property*, New York: Harcourt, Brace & World.

Biondi, Yuri (2002), *Gino Zappa e la rivoluzione del reddito. Azienda, moneta econtabilità nella nascente Economia aziendale* [Gino Zappa and the business income to the firm. Firm, money and accounting in the Italian business economics revolution], Padova: CEDAM.

—— (2003), "*La nature économique de l'entreprise au croisement des théories économiques et de la comptabilité*" [Economic organization between economic theories and accounting], unpublished PhD thesis, December 5, Universities of Brescia and Lyon 2.

—— (2005a), "The Firm as an Entity: Management, Organization, Accounting," *Paper (Economia aziendale Dept.)*, no. 46, University of Brescia, 2 August 2005 in the referenece Biondi (2005a).. Available online at http://ssrn.com/abstract=774764 (accessed 20 January 2006).

—— (2005b), "*G. Zappa, T. Veblen, J.R. Commons e l'impresa come istituzione economica*" [Zappa, Veblen, Commons on the firm as an economic institution], *Storia del Pensiero Economico*, (1): 93–120.

—— (2006), "Accounting and the economic Analysis of the Firm as an Entity," in M. Dietrich (ed.), *The Economics of the Firm: Analysis, Evolution and History*, New York and London: Routledge.

Blair, Margaret (2003), "Locking in Capital: What Corporate Law Achieved for Business Organizers in the Nineteenth Century," *UCLA Law Review*, 51 (2): 387–455. Available online at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=495984 (accessed 15 September 2006).

Buchanan, James M. and Thirlby, F.G. (eds) (1981), *L.S.E. Essays on Cost*, New York and London: NY University Press (1st edn, London: LSE, 1973).

Coase, H. Ronald (1937), "The Nature of the Firm," *Economica*, new series, no. 4, November: 386–405.

—— (1960), "The Problem of Social Cost," *Journal of Law and Economics*, 3 (1): 1–44.

—— (1981), "Business Organization and the Accountant" [shortened version of a series of twelve articles from *The Accountant*, 1.10.1938 – 17.12.1938], in J.M. Buchanan, G.F. Thirlby (eds), *L.S.E. Essays on Cost*, New York and London: NY University Press.

—— (1988 [1991]), "The nature of the firm: Origin, Meaning, Influence," *Journal of Law, Economics, and Organization*, 4 (Spring): 3–47. also in O.E. Williamson and S.G. Winter (eds) *The Nature of the Firm. Origins, Evolution, and Development*, Oxford: Oxford University Press.

—— (1990), "Accounting and the theory of the firm," *Journal of Accounting and Economics*, no. 12, 1990: 3–13, reprinted in this volume.

Commons, R. John (1924), *Legal foundations of capitalism*, New York: Macmillan; reprinted 1996, USA: Transactions Pub.

—— (1934), *Institutional economics. Its place in political economy*, New York: Macmillan; reprinted 1990, USA: Transactions Pub.

Copeland, A. Morris (1927), "An Instrumental View of Part–Whole Relation," *The Journal of Philosophy*, 24 (4), February: 96–104.

Crémer, Jacques (1986), "Cooperation in Ongoing Organizations," *Quarterly Journal of Economics*, CI (1), February: 33–50.

Demsetz, Harold (1995), *The economics of the business firm, Seven critical commentaires*, Cambridge: Cambridge University Press.

Favereau, Olivier (1997), "L'incomplétude n'est pas le problème, c'est la solution" [Incompleteness is not the problem, but the solution], in B. Reynaud (ed), *Les limites de la rationalité. Tome 2: les figures du collectif*, Paris: La Découverte.

Friedman, Milton (1953), *Essays in positive economics*, Chicago: University of Chicago Press.

Grossmann, Sandford and Hart, Oliver (1986), "The costs and benefits of ownership: a theory of vertical and lateral integration," *Journal of Political Economy*, 94: 691–719.

Gruchy, G. Allen (1947), *Modern Economic Thought*, New York: Prentice-Hall Inc.

Gynther, S. Reginald (1967), "Accounting Concepts and Behavioral Hypotheses," *The Accounting Review*, XLII (2), April: 274–90.

Hart, Oliver (1989), "An Economist's Perspective on the Theory of the Firm," *Columbia Law Review*, 89 (7): 1757–74. reprinted in O.E. Williamson (ed.) (1990), *Organization and theory from C. Barnard to the present and beyond*, Oxford: Oxford University Press.

Hodgson, M. Geoffrey (2002), "The Legal Nature of the Firm and the Myth of the Firm-Market Hybrid," *International Journal of the Economics of Business*, 9 (1): 37–60.

Hölmstrom, Ben (1999), "The firm as a sub-economy," *Journal of Law, Economics and Organization*, 15: 74–102.

Ijiri, Yuji (1967), *The Foundations of Accounting Measurement: A Mathematical, Economic and Behavioral Inquiry*, Englewood Cliffs: Prentice-Hall.

—— (1975), *Theory of Accounting Measurement*, Studies in Accounting Research, no. 10, USA: American Accounting Association.

Kirat, Thierry and Bazzoli, Laure (1998), "*L'entreprise et les règles juridiques: une perspective d'économie institutionnaliste*" [The firm and the legal rules: an institutional economic perspective], *Revue interdisciplinaire d'études juridiques (RIEJ)*, 40: 47–78.

Kirman, Alan (1987), "The intrinsic Limits of modern economic Theory. The Emperor has no Clothes," EUI working paper no. 87/323, Department of Economics, Florence: European University Institute.

—— (1997), "The economy as an evolving network," *Journal of Evolutionary Economics*, 7: 339–53.

Littleton, A. Charles (1928), "What is Profit?," *The Accounting Review*, III (3), June: 278–88.

—— (1938), "A substitute for stated capital," *Harvard Business Review*, 17 (1), Autumn: 75–84.

—— (1953), *Structure of Accounting Theory*, USA: American Accounting Association.

—— (1961), *Essays on Accountancy*, C.A. Moyer (ed.), Urbana: University of Illinois Press.

Machlup, Fritz (1946), "Marginal Analysis and Empirical Research," *American Economic Review*, 36 (4), part I, September: 519–54.

—— (1967), "Theories of the firm: marginalist, behavioral, managerial," *American Economic Review*, LVII (1), March: 1–33.

Pareto, Vilfredo (1906), *Manuale di Economia Politica, with Proemio: "Introduzione alla scienza sociale*" [Manual of Political Economy, with Preface: "Introduction to the social science"], Padova: CEDAM, 1974. First French Edition: Lausanne 1909, in V. Pareto (1981), *Œuvres Complètes*, Tome VII, Geneva: Librairie Droz.

Perroux, François (1966), "L'entreprise et le phénomène collectif du second XXe siècle" [The firm and the collective phenomenon of the late XXth century], I, 1–17; in F. Bloch-Lainé and F. Perroux (eds), (1966), *L'entreprise et l'économie du XXo siècle. Étude internationale* [The Firm and the Economics of the XXth Century. An International Inquiry], Three vols, Paris: PUF.

Raby, L. William (1959), "The Two Faces of Accounting," *The Accounting Review*, 34 (3), July: 452–61.

Rajan, Raghuram G. and Zingales, Luigi (1998), "Power in a Theory of the Firm," *Quarterly Journal of Economics*, 113 (2): 387–432.

—— (2001), "The Firm as a Dedicated Hierarchy: A Theory of the Origins and the Growth of Firms," *Quarterly Journal of Economics*, 116 (3): 805–51.

Royce, Josiah (1914), "Principles of Logic," in A. Ruge (ed.), *Encyclopedia of the Philosophical Sciences*, I – Logic, London: Macmillan.

Ruben, David-Hillel (1983), "Social Wholes and Parts," *Mind*, 92 (366): 219–38.

Schumpeter, Joseph A. (1912), *Theorie der wirtschaftlichen Entwicklung* [The Theory of economic Development], Leipzig: Duncker&Humblot; 2nd edn: 1926.

Scott, A. Richard (1979), "Owners' Equity, The Anachronistic Element," *The Accounting Review*, 54 (4), October: 750–63.

Shubik, Martin (1993), "Accounting and its relationship to general equilibrium theory," in M. Shubik (ed.), *Accounting and Economics, Economic Notes* by Monte dei Paschi di Siena, 22 (2): 226–34; reprinted in this volume.

Simon, A. Herbert (1991), "Organizations and markets," *Journal of Economic Perspectives*, 5 (2), Spring: 25–44; reprinted in this volume.

—— (1997), *An empirically based microeconomics*, Cambridge: Cambridge University Press.

Stauss, James H. (1944), "The Entrepreneur: The Firm," *Journal of Political Economy*, 52 (2), June: 112–27; reprinted in this volume.

Sprouse, T. Robert (1957), "The Significance of the Concept of the Corporation in Accounting Analyses," *The Accounting Review*, XXXII (3), July: 369–78.

Veblen, Thornstein (1900), "Industrial and Pecuniary Employments," Publications of the American Economic Association (AEA), series 3, vol. II; in T. Veblen (1919), *The Place of Science in Modern Civilisation and other Essays*; reprinted 1969, New York: Capricorn.

—— (1904), *The Theory of Business Enterprise*, New York: Scribner's; reprinted 1978, USA: Transaction Books.

—— (1923), *Absentee Ownership*, New York: B.W. Huebsch; reprinted 1954.

Whitehead, N. Alfred (1932), "Objects and Subjects," *The Philosophical Review*, 41 (2), March: 130–46.

Williamson, E. Oliver (1975), *Markets and Hierarchies. Analysis and Antitrust Implications*, New York: Free Press.

—— (1988), "Corporate Finance and Corporate Governance," *Journal of Finance*, 43 (3), July: 567–91.

—— (1991), "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives," *Administrative Sciences Quarterly*, 36, June: 269–96.

Williamson, E. Oliver and Winter, G. Sidney (eds) (1991), *The Nature of the Firm. Origins, Evolution, and Development*, Oxford: Oxford University Press.

Zappa, Gino (1927), *Tendenze nuove negli studi di ragioneria* [New Trends in Accounting], inauguration speech for the academic year 1926–7 at R. Istituto Superiore di Scienze Economiche e Commerciali di Venezia on 13 November 1926, Milan: IES.

—— (1937), *Il reddito d'impresa* [The Income to the Firm], Milan: Giuffrè; 1st complete edn 1929, Rome: ALI.

Zingales, Luigi (2000), "In Search of New Foundations," *Journal of Finance*, 55 (4): 1623–53.

# 15 Some building blocks for a theory of the firm as a real entity

*David Gindis*

## Introduction

The firm is a real entity and not an imaginary, fictitious or linguistic entity. This implies that the firm as a whole exhibits a sufficient degree of unity or cohesiveness and is durable and persistent through time. The firm is essentially composed of a particular combination of constituents that are bound together by something that acts as an "ontological glue," and is therefore non-reducible to other more basic entities, i.e., to its parts or its members. From our perspective, the firm is not simply an aggregate or a collection. It is a real integrated entity and a dynamic causal system. Institutional and organisational aspects enter the picture. These assertions stand in sharp contrast with mainstream theories of the firm whose proponents are more preoccupied with questions of contractual provisions, vertical integration or opportunism than the general and more fundamental questions related to what firms really are.

The dominant contractual paradigm establishes the nature of the firm as a "nexus of contracts" (Jensen and Meckling 1976) or a "collection of assets" (Hart and Moore 1990), and further questioning of these constructs is left out of most, if not all, mainstream papers on the theory of the firm. When Cheung (1983: 3) claims that "we do not know what the firm is – nor is it vital to know," we are left with a simple "shorthand description" of a set of contractual relations. When Jensen and Meckling (1976: 311) brush off the question by stating that the firm is simply a "legal fiction," we are left with a "word" with as little substantive content as words referring to imaginary creatures. It is the object of this chapter to provide the theoretical grounds for rejecting these views. Firms are real entities that need to be theoretically treated as such.

The rest of the chapter is organised as follows. The second section establishes the main tenets and variations of the dominant theories of the firm before discussing some central problems related to the underlying ontological status of the firm. We claim that these views are so strong in denying the existence of the firm that they can hardly be said to be theories of the firm. Given that these theories are modern revivals of some old views, we revisit the turn-of-the-century debates in the third section in order to reconsider the merits of various alternative entity theories, adopting an analytical rather than historical perspective.

Building on these insights, the fourth section provides some building blocks for a theory of the firm as a real entity. We argue that such a theory needs to overcome the impasse of reductionism and account for the firm's cohesiveness and durability through time. In the final section we discuss some implications of our analysis for the theory of the firm, and apply some of our arguments to multi-unit structures such as corporate groups and vertical forms of network organisation.

## The dominant theories of the firm in economic theory

Our interest in this section lies mainly in synthesising the "nexus of contracts" and "collection of assets" views of the firm, that are without a doubt the dominant views in economic theory. Arguably, the nexus of contracts view is also strongly established in corporate law theory. We therefore ignore other mainstream theories of the firm. We examine the underlying ontological status of the firm in the nexus of contracts and collection of assets views, and argue that both conceptions lack a meaningful theory of the nature of the firm. Their ontological individualism implies that anything other than individuals, their contractual relations and their assets is merely fantasy. This assessment paves the way for an alternative theory of the firm discussed in the third and fourth sections of the chapter.

### *The firm as a nexus of contracts*

The modern nexus of contracts theory of the firm goes back to Alchian and Demsetz (1972: 794), according to whom "the essence of the classical firm is . . . a contractual structure." Picking up on this insight, Jensen and Meckling (1976: 310) agree that "contractual relations are the essence of the firm." In short, when "the firm is viewed as a set of contracts" (Fama 1980: 289), it is not "vital to know" what the firm is, since "the word 'firm' is simply a shorthand description of a way to organize activities under contractual arrangements" (Cheung 1983: 3). In other words, "what is called a firm is a special set of contracts among owners of resources used in the coalition" (Alchian 1984: 34). Clearly, if the firm is seen as a "set of contracts," then the firm must also be seen as a set of individuals entering those explicit and implicit contracts.

The firm is thus conceptualised both as a set of contracts and as a "coalition" or an "association" between resource owners working more or less as a team. In fact, teamwork and contractual coalition arguably boil down to the same thing: "Does the essence of the 'firm' lie in teamwork or in the nexus of long-term contracts (i.e., agreements restraining the behavior of transactors)? . . . Teamwork always involves such contracts. We . . . can think of neither significant nor interesting cases where teamwork does not create dependencies calling for contractual restraints" (Alchian and Woodward 1988: 70). Given that contracts restrain opportunistic individuals, contracts are the essential feature of the firm. Of course, given the ubiquity of the contract, other forms of association as well as market transactions are said to be contractual.

Thus, contrary to Coase's (1937) theory of the firm–market dichotomy and Williamson's (1991) attempt to analyse "discrete forms," Jensen and Meckling (1976: 311) state that "it makes little or no sense to try to distinguish those things that are 'inside' the firm . . . from those things that are 'outside' of it." This makes sense from the nexus of contracts view, insofar as the world is regarded as a contractual continuum. Forms of economic organisation differ only in degree since they all share the same contractual essence. Indeed, the contractual theory "makes the boundary of the firm fuzzy: a bright line distinguishing 'inside' and 'outside' is missing" (Alchian and Woodward 1988: 76). In this conception of the world, "aspects of firm-like contractual arrangements brush aside the question of absolutes – 'When is a nexus of contracts *a firm*?' – and substitute instead a question of relatives – 'When is a nexus of contracts *more firm-like*?'" (Demsetz 1988: 155, emphasis in original).

Jensen and Meckling (1976: 310–11) stress that "it is important to recognize that most organizations are simply *legal fictions which serve as a nexus for a set of contracting relationships among individuals* . . . The private corporation or firm is simply one form of *legal fiction*" (emphasis in original). Thus, when contractarians claim that the firm is the "common signatory of a group of contracts" (Hansmann 1996: 18), they are referring to nothing more than a convenient "legal fiction," understood as "the artificial construct under the law which allows certain organizations to be treated as individuals" (Jensen and Meckling 1976: 310, n.12). But in no sense should the firm be really viewed as such. Only individuals can have objective functions, can own and invest in assets, negotiate and sign contracts. It makes no sense to speak of the "behaviour" of firms, since firms are not individuals and only individuals can act, whether responsibly or not.[1] It should be clear that "the 'personhood' of a corporation is a matter of convenience rather than reality" (Easterbrook and Fischel 1991: 11).

In this context, Easterbrook and Fischel's (1985: 89) much-cited discussion of limited liability reminds us that "the liability of 'the corporation' is limited by the fact that the corporation is not real. It is no more than a name for a complex set of contracts among managers, workers, and contributors of capital. It has no existence independent of these relations." Whatever the expression used, i.e., "firm," "corporation," "nexus of contracts," or "legal person," for contractarians these are no more than shorthand ways of referring to individuals and their relations. Whether an "association" or a "coalition," the firm is the sum or aggregate of its individual members. In this sense, "it is *not* silly to consider the entry of a new stockholder to be the creation of a new firm" (Alchian 1984: 47, emphasis in original). In dissolving the firm and dismissing it as a fiction, nexus of contracts theorists mean that shareholders own the corporate assets.[2]

## The firm as a collection of non-human assets

We now turn to the second mainstream conception of the firm, according to which the firm is a collection of non-human assets. Our presentation of this

approach is quite brief because the views of Grossman, Hart and Moore are simply stated when one has in mind the nexus of contracts point of view. Bluntly put, Grossman and Hart (1986: 693) "define a firm to consist of those assets that it owns or over which it has control." In this spirit, when looking for the firm in the web of contractual relations, the rule is simple: "identify a firm with the assets it possesses" (Hart and Moore 1990: 1120). This property rights view of the firm excludes human assets or human capital, since human assets are inalienable, i.e., they cannot be bought or sold.

Such a statement of the nature of the firm as a collection of non-human assets contrasts with the nexus of contracts view, which "does less to resolve the questions of what a firm is than to shift the terms of the debate" (Hart 1989: 1764). Here, the boundaries question is intimately linked to the nature of the firm question. Indeed, "one can . . . sidestep the issue entirely, by arguing that everything is contractual, and that firms are a mirage [i.e.,] they are simply 'standard-form' contracts . . . But if firms are a mirage, it is difficult to explain the enormous resources that firms expend merging and breaking up" (Moore 1992: 494). Thus, property rights are essential to the theory of the firm since they provide an account of firm boundaries: assets that belong to the firm are "inside" firm boundaries and assets that do not are "outside."

In Hart and Moore's (1990) analysis, agents form coalitions depending on the expected value of their participation, which in turn depends on the effects of asset ownership. However, whereas in the nexus of contracts view the firm is the central contracting party and thus ensures the connection between the asset owners composing the coalition, in the theory of the firm as a collection of assets, the link is different. Since "*a firm is identified with the collection of physical assets* over which the owner . . . has the residual control rights" (Moore 1992: 496, emphasis in original), the link is the concentration of property rights in the hands of one agent. The coalition of owners disappears and we are left with "*the* owner" that personifies the aggregate of owners. What, then, is the link between the owner and the other agents necessary for production? The existence of such a link is a critical aspect of a theory of the firm, and Hart (1995: 57) rightly underlines that "without something to hold the firm together, the firm is just a phantom." According to the theory of the firm as a collection of assets, the answer is obvious:

> A firm's nonhuman assets . . . simply represent the glue that keeps the firm together . . . If such assets do not exist, then it is not clear what keeps the firm together . . . One would expect firms without at least some significant nonhuman assets to be flimsy and unstable entities, constantly subject to the possibility of break-up or dissolution. My impression is that the (causal) evidence is not inconsistent with this view.
>
> (Hart 1995: 57–9)

The owner exercises control over nonhuman assets and this leads to an indirect control of human assets. The glue that ensures the link between human

assets and the nonhuman assets, i.e., the firm, is thus one based on the exclusion and exit rights of the assets' owner: employers have considerable leverage by virtue of owning the nonhuman assets that workers need in order to be productive. We do not further develop the logic of the model, since our interest lies in the nature of the firm. We pick up on the matter of the "glue" below.

### *Underlying ontological commitments*

The preceding theories of the firm present some serious theoretical problems. Indeed, both views lack an adequate theory of the nature of the firm. In what follows, we make apparent the ontological commitments of these theories: either they make an outright denial of the existence of the firm, or they explain away the firm completely by concentrating on individuals. Commenting on the widespread adoption of the nexus of contracts view, Bratton (1989a: 409) writes: "some have accorded this notion the weight of scientific truth: It has been received in the legal literature as an ontological discovery with immediate and significant implications for corporate law discourse." Ironically, the "ontological discovery" is that the firm is a "fiction," i.e., that it does not really exist.

Jensen and Meckling (1976: 311) admit that the firm in their presentation has "little substantive content." There is no point in marking out "firm boundaries," since the concept itself makes little or no sense in a world that is a contractual continuum. The only possible difference between observed forms of organization is a difference in degree. There can be no differences in nature if the essence of everything is identical. The "firm" can be no more than the by-product of the study of contractual relations. This sort of reasoning is straightforward in denying the firm ontological status, although it relies on notions of "firm-likeness" which beg a definition of the firm that should logically pre-exist. Fundamentally, the whole cannot be significantly different from the sum or aggregate of its parts. Indeed, if the firm does not exist apart in any way from its constituent parts, be these individuals, contracts and/or assets, then the issue of comparing the whole with the sum of the parts makes little sense.

A further problem of intellectual coherence can also be raised. If the world is a contractual continuum, which of the existing or possible contracts are those that define or belong to a particular firm? The reasoning cannot escape circularity, since one cannot identify a nexus of contracts independently of a given firm. Surely a firm cannot consist of all those contracts that are linked to the firm, because that "would be like saying that a zebra is a nexus of stripes linked to a zebra" (Eisenberg 1999: 830). The same is logically true of the collection of assets view of the firm that holds that a zebra is a collection of stripes owned by a zebra. Crucially, there is no essential difference, or difference in nature, between an individual owner of one asset and an individual owner of several assets, or between one asset and a set of assets. All individuals are owners of some assets. Are all individuals "firms"? Likewise, is any non-human asset a "firm"? To use Eisenberg's words again, it would be like saying that a zebra is a collection of stripes and that a one stripe is a zebra.

Can a "collection" of X count as the "nature" of Y? Hart (1995) needs a theory of something that can act as the "glue," because without it the firm is "just a phantom." Hart is correct. Without the glue, a "collection of assets" is no different from a heap of sand easily blown away on a windy day. Heaps of sand composed of random grains of sand are indeed "flimsy and unstable entities, constantly subject to the possibility of break-up or dissolution" (Hart 1995: 58–9). Yet Hart's idea of the glue is poor. If "a firm's non-human assets . . . simply represent the glue that keeps the firm together" (Hart 1995: 57), then the answer to the question "what is a firm?" is "a collection of non-human assets" and that the answer to the question "what holds the firm together?" is "a collection of non-human assets." Regardless of this logical problem, surely, the threat of exclusion alone cannot account for whatever holds the firm together and points more to the (potential) dissolution of the firm than to its (actual) unity. Far from being the sort of thing that could bind anything together, a collection is itself in need of being bound together if it is to form a whole.

Our critical assessment of the nexus of contracts and collection of assets stories thus reveals that these are in no meaningful sense theories of the nature of the firm. Therefore, both views cannot be said to be theories of the firm. Both are theories of "collections" of some sort and not of "firms." Words such as "firm" may be used in order to facilitate discussions with the layman or even with the specialist, but the word "firm" refers to sets of individuals, contracts and/or assets and is simply used for convenience. Firms exist only linguistically or conceptually. Firms are dissolved and dismissed as fictions. Contractual theorists in general and those cited in particular are ontological individualists. Anything can and everything should be reduced to individuals.

Firms do not act. Firms do not have intentions or objective functions. "Their "behavior' is like the behavior of the market . . . We seldom fall into the trap of characterizing the wheat or stock market as an individual, but we often make this error by thinking about organizations as if they were persons with motivations and intentions" (Jensen and Meckling 1976: 50). This comparison with the impersonally efficient market is supposed to make economists think twice before ever again committing the error of thinking that the firm is an individual. However, to say that the firm is an "individual" is different from saying that it is a "person" in that the latter is a moral and legal term. Corporate personality may indeed be a legal fiction but individuality implies that firms may be said to distinctly exist and act in a unified sort of way – not that the corporation is an individual. This line of inquiry is important for our discussion throughout the rest of this chapter, in which we seek to build an alternative theory of the firm as a real and not fictitious or linguistic entity.

## From fictions and aggregates to real entities

We have argued that the nexus of contracts and collection of assets theories fail as theories of the nature of the firm. Given that both views are modern revivals of some old debates that have been evolving in a cyclical fashion for many

centuries (Avi-Yonah and Sivan, this volume), a number of contemporary critics of the current mainstream have called for the revival of the entity theory of the turn of the twentieth century (e.g., Blair 1999; Smith 2001).[3] Reviving real entity theory is also our task in this section. Keeping in mind that "history offers ideas and approaches that are better and worse, and [that] the most recent is not always the best" (Smith 2001: 72), instead of tackling the debate from a general historical perspective, we adopt an analytical perspective that seeks to disentangle the ontological from the normative aspects involved. We first underline the links between fiction, aggregate and property conceptions of the corporation, before exploring the alternative entity view.[4]

### *Fictions and aggregates*

In most jurisdictions in the world, the corporation is a "legal person," a "legal entity" separate from the various natural persons composing it or having an interest in its economic activities. Legal personality means that the corporation can own assets, sign contracts, sue or be sued, and so forth. However, under the doctrine of "piercing the corporate veil," courts may decide to break the legal protection offered by personality if the mask of separate personality is used fraudulently. In such cases, the law treats the corporation as a fiction and makes the natural persons involved liable for corporate torts or debts. This doctrine seems to imply that if we pierce the corporate veil, then we find nothing but real individuals. Many interpret this to mean that there was no real entity to begin with. While these issues have traditionally been applied to "corporations," they have been transposed to other types of "companies" by recent legal evolutions.[5] Accordingly, although in this section we discuss the corporation following traditional debates, our arguments apply more generally to all forms of companies, i.e., to the firm.

As is well known, the "fiction theory of the corporation" is as old as Roman law and was often wielded by medieval canonists and civilists. The corporation was generally considered as an imaginary or legal being that could be nothing more than the individuals composing it. That this conception is closely linked to the "aggregate theory of the corporation" is quite obvious: since the corporation is nothing more than the aggregate of its real individual members, all reference to the collective corporate entity is reference to a "legal fiction" that is used merely for convenience. Proponents of this approach believed that "the rights and duties of an incorporated association are in reality the rights and duties of the persons who compose it, and not of an imaginary being" (Morawetz 1882: 2). Jensen and Meckling (1976), and others, use both the fiction theory and the aggregate theory. Of the various possible aggregates, the aggregate of shareholders or owners may be called the "proprietary theory of the corporation." Hart's (1995) collection of assets view based on property rights is clearly a variant of this aggregate theory. In dissolving the firm and dismissing it as a fiction, contractarians declare that the aggregate of shareholders is the owner of the aggregate of corporate assets.

These views carry direct implications for the much-debated question of how firms should be run. Of course, following Jensen and Meckling (1976), Butler and Ribstein (1989) and Easterbrook and Fischel (1991), in a private property market economy firms should be run in the exclusive or at least the predominant interest of shareholders. The only possible answer to Dodd's (1932) famous "For Whom Are Corporate Managers Trustees?" question is "The Shareholders." Again, in dissolving the firm and dismissing it as a fiction, the basic agency model examines direct relations between the aggregate of shareholders and firm managers, the latter simply managing the former's assets. A popular rationale for this "shareholder primacy norm" is based on the belief that the category "shareholders" exhibits more homogeneous interests than any other "stakeholders" (e.g., Hansmann 1996) and that shareholders, as residual claimants, are also the main risk-bearers in the undertaking of economic activities.

However, Stauss (1944, reprinted in this volume) correctly criticised the mainstream myth of the existence of some unique class of individuals that are "homogenous risk-bearers" when in fact it is the firm as a whole that undertakes risky entrepreneurial activities. Arguably, shareholders, as defined by Hansmann, can themselves be considered "fictional," a category that the proprietary conception of the corporation needs in order to function. In fact, "shareholders . . . are a legal fiction, and in many ways a far more problematic fiction than the corporation itself" (Greenwood 1996: 1025). If the law is replete with fictions, what makes any one of them so special? The underlying ontological individualism of the dominant theory explains why shareholders are considered "more real" than corporations, since corporations are said to exist only linguistically or conceptually.[6]

Many of those who today invoke the entity theory of some hundred years ago do so with the intent to influence the current corporate governance debate. Thus, entity sympathisers often remind us that managers legally owe fiduciary duty to the corporation itself rather than to the shareholders, that it is the corporation as a legal entity that owns corporate assets rather than the shareholders directly, and that no one owns the corporation (e.g., Blair 1995; Greenwood 1996; Iwai 1999; Ireland 1999; Stout 2005). Alongside these explicit legal features against the dismissal of corporations as fictions, the more general "stakeholder theory of the corporation" (e.g., Donaldson and Preston 1995) reminds us that individuals and groups of individuals other than shareholders are also part of the economic nexus, and that running the firm in the sole interests of shareholder value is insufficient. In its crude version, stakeholder theory is a form of aggregate theory that simply denies the supremacy of the proprietary conception. In a more elaborate version, stakeholder theory views the corporation as a "social entity" or institution having social responsibilities.

Hence the general position of many proponents of corporate social responsibility today. Blair (1995) thus opposes two families of theories of the corporation, namely the "property theory of the corporation" which underlies the current shareholder primacy norm and the "social entity theory of the corporation." Although we share Blair's (1999) view that the entity theories of some hundred

years ago need to be revived, we depart from the corporate governance and corporate social responsibility debates and seek analytical insights for a theory of the firm only.[7] We need to disentangle the ontological from the normative aspects of the real entity theory. We agree with Smith (2001: 70), who believes that "in the decades ahead, economists and legal scholars will become more sophisticated in their appreciation of human groups and of social forms such as corporations." This is precisely why we seek to provide some building blocks for a theory of the firm as a real entity. In what follows, we suggest the revival of entity theory through Freund's (1897) major but neglected work.

## *Reviving real entity theory*

The aggregate and fiction theories of the corporation of a hundred years ago were much criticised by authors influenced by Continental theories of "real corporate personality," associated with Hegelian legal theorists such as Gierke. Expressions such as "the personality of the corporation . . . is in no sense . . . artificial or fictitious, but is every whit as real and natural as is the personality of man" (Maitland 1900: 335), or "a corporation is an entity – not imaginary or fictitious, but real, not artificial but natural" (Machen 1911: 262) became quite common. As opposed to the fiction theory, the "real entity theory of the corporation," also called the "natural entity theory of the corporation," holds that the corporation not only exists separately from its members but also can literally be said to act and have "volition." As opposed to the proprietary theory, the corporation's being a "person" implies that it cannot be considered an object of property rights.

Some entity theorists had larger preoccupations than the specific legal form of the corporation, extending their analysis to unincorporated bodies (Maitland 1903) or to associations and partnerships, all seen as "more than the aggregate of [their] members" (Rowley 1931: 560). From this point of view, the firm's existence has nothing to do with the law. The law simply complies with the fact of its existence and attributes legal capacity, i.e., recognition inside a particular legal system, to an already existing economic capacity, i.e., regardless of the legal system. Machen (1911: 261) thus states: "A corporation exists as an objectively real entity . . . The law merely recognizes and gives legal effect to the existence of this entity. To confound legal recognition of existing facts with creation of facts is an error." In the same spirit, Laski (1916: 422) claims that "the entities the law must recognise are those which act as such, for to act in unified fashion is – formality apart – to act as a corporation." In this sense, this movement was in search of a realistic approach regardless of legal form.[8] Berle's (1947, reprinted in this volume) theory of "enterprise entity" as different from "corporate entity" is an important later contribution to this approach.

In order to revive the entity theory, it is necessary to isolate these ontological aspects from the ever-present issue of corporate personality. Many entity theorists conflated "reality" and "naturalness," and accorded excessive weight to the "personality" of this natural entity. In fact, most debates and most of the

confusion arose and still arise from the thorny issue of corporate personality. Therefore, given that corporate existence and corporate personality are all too often equated, this is the first issue that needs to be dealt with. We need to go beyond the fallacious assumption that if a collective entity is said to be treated in legal terms as a "legal person" this means that the collective entity *is* a person (Hodgson 2002). Recall that Jensen and Meckling (1976) dismiss the firm as a fiction on the grounds that its personality is a fiction. Yet legal fictions need not be ontological fictions (Dejnožka, 2006). It is also important to note that what the law treats as a "person" does not necessarily or effectively correspond to a "human being" or an "individual."

Fundamentally, the term "person" derives from the Latin *persona* and the Greek *prosopa* meaning "mask" in a dramatic representation. More precisely, "person" can refer to the mask worn by an actor, the actor behind the mask or the actor-in-the-mask. In his discussion of Vining's (1978) theory of "legal identity," Dejnožka (2006: 30) points out that "courts treat legal persons as human-beings-in-persons, corresponding to actors-in-masks. If [so], the legal persons are more than masks or roles. They are human beings in masks or roles. And a corporate person is a group of human beings in a collective mask or role." Such a collective mask is a social institution and this can be said without moral or normative intentions or preconceptions. What is behind the mask of legal identity is a real complex entity. This holds for all legal forms of business companies.

Reviving entity theory also involves steering away from ideas of "organisms" and "groups souls" or "vital forces" that various entity theorists pronounced. It should be equally clear that our task is not to establish the "ultimate moral unit" sought by Maitland (1905) and many others. Rather, closer in spirit to Laski and Rowley, we need a causally grounded theory and not a morally grounded theory. Ontology is independent from and logically precedes the normative question. We are looking for a theory of the firm as "a relatively coherent and stable whole," since this is in effect our working definition of a real entity. In this context, Freund (1897) has had surprisingly few commentators, considering that his theory rejects the fiction and aggregate views of the corporation while avoiding the pitfalls of organicism and does not share the normative bent of other writers of his time.

Although speaking from the legalistic perspective, in an actually ontological and therefore morally neutral spirit, Freund (1897: 47) lists three "salient characteristics of the body corporate: its unity, its distinctiveness and its identity in succession." All three derive from a "representation principle" that is needed, according to Freund, in order to acknowledge the existence of higher-level rights and agency as opposed to individual rights and agency. In a nutshell, the representation principle is that corporate organization and corporate rules bind individual agency in such a way that one can properly speak of corporate unity, distinctiveness and retention of identity through change. For Freund, if these salient features are in fact present in a given association, one can speak of a real entity. The difficulty is to show how common purpose and collective action produce a level of unity, distinctiveness and identity sufficient for a corporation

to be a real entity without appealing to any literally volitional or moral features. Freund says:

> The association becomes visible and active in and through individuals only, but the common purpose, the concerted action, and the combined resources, produce upon our mind the impression, that the association itself enjoys something like the power of individual personal agency. The resulting conception is not one of absolute unity, such as the German jurists demand . . . but a relative unity . . . There is no absolute objective test by which we could be forced to allow or deny the character of unity to an aggregate body of human persons. The analogy of composite things explains perfectly the nature of the association.
>
> (Freund 1897: 77)

Freund rightly insists on the unity of complex wholes composed of human and non-human integrated parts. His analysis analogises corporations to composite artefacts. His theory of the reality of the whole as not literally an agent but a causally interacting composite of agents and non-agents is consistent with the (literally false) phenomenological impression that the association itself enjoys the power of agency. Freund's three features of the corporation are what Rowley (1931) and much later Khalil (1997) called "individuality." We prefer Freund's original terms. It is important to understand that this composite unity may be imperfect at times and that this is something that the law may or may not capture. The group's organisation may or may not have the relative unity or wholeness that allows one to properly speak of collective action.

The form and exercise of intra-association control are both essential elements of a theory of collective action. But while the power to act as a group is an essential element of the whole's reality, there is no need to suppose that there exists some sort of supra-individual or metaphysical "soul" or "vital force" of the corporation. In sum, Freund's (1897) approach paves the way for our analysis below which takes further theoretical steps. Stauss's (1944: 112, reprinted in this volume) subtle position is equally relevant as a point of departure for our own theory: "the entity known . . . as *the firm* is taken as a real institution. As such the firm *exists* apart from the individuals who compose its decision-making organization, but it *does not function* apart from them. Thus the entity is not a fiction; it is a fact" (emphasis in original). Generally speaking, "the entity commonly known as 'corporate entity' takes its being from the reality of the underlying enterprise, formed or in formation" (Berle 1947: 344, reprinted in this volume). This applies to all legal forms of business companies.

## Building blocks for a theory of the firm as a real entity

Based on our preceding revival of entity theory through Freund's (1897) important work, our present task is to further develop a theory of the firm as a real entity. In the spirit of Weissman (2000), Mäki (2001) and Bunge (2003), we

believe that economic ontology must include firms as fundamental wholes non-reducible to other more basic entities, i.e., to their parts or members. We argue that various forms of reductionism are untenable when theorising firms and other complex human organisations. The firm as a real complex entity is at least as real as its members (Dejnožka 2006). From an ontological point of view, the key issues are parts–whole relations, unity, wholeness, cohesiveness, persistence, durability, identity and distinctiveness. From an economic point of view, the directly related issues include collective action, institutions, organisation, managed economic system, competence, knowledge stickiness and heterogeneity. From a legal point of view, the issues are substance over form, or real enterprise over legal entity. In arguing against reductionism we provide a theory of the "ontological glue" that accounts for the firm's cohesiveness and durability through time.

## *Beyond reductionism*

Firms are not simply sets of contracts or collections of assets. The firm is not an epiphenomenon. To reduce the firm to its constituent parts "is no more reasonable than is treating a human being as no more than the chemicals that make her up" (Greenwood 2005: 15). Although mereological relations, i.e., complex parts–whole processes, need to be part of a theory of the firm, we reject "mereological essentialism" (Chisholm 1973), that holds that parts are essential to their wholes in the sense that the whole is determined by the parts. This is a typically reductionist position since the whole does not exist independently of its parts and any modification of the parts implies a modification of the whole. Clearly, the aggregate theories discussed throughout this chapter, and in particular Alchian's (1984: 47, emphasis in original) view that "it is *not* silly to consider the entry of a new stockholder to be the creation of a new firm," are strictly speaking compatible with mereological essentialism. This presents a conceptual problem for the identity or durability of the whole through time.

Firms are not aggregates or "mereological sums." The very issue of comparing the whole to the sum of its parts is problematic. In his discussion of the problem, Copeland (1927) rightly stresses that for parts or whatever to be summed together they need to be of the same magnitude. For example, how does one "add" individuals and non-human assets or artefacts? Addressing this question, Ruben (1983) argues that individuals cannot be said to be "parts" of social wholes since "being a part of" is a purely mereological relation whereas when we think of individuals as "members" of such wholes we have in mind a particular social relation not reducible to mereological relations. While we can think of non-human assets or artefacts as "parts" of firms, we need to admit the theoretical superiority of individuals as "members" of firms or other organisations since this connotes the members' agential power through which firms function. In any case, the whole is reducible neither to its members (e.g., owners) nor to its parts (e.g., assets).

The fact that reductionism fails as an adequate analysis of the nature of the firm implies that if one systematically reduces the firm as a whole to anything else one necessarily loses sight of a great deal of the picture. It is important to stress that firms as structured wholes compete with other firms as structured wholes, that firms as wholes can sue and be sued (even though individuals go to court as actors-in-the-mask), can produce and be competitive. Firms as wholes possess capabilities. Firms as wholes have temporal reputations in transactions and generate income or suffer losses. Firms as wholes undertake risk-bearing activities (Stauss 1944, reprinted in this volume). There is nothing imaginary or fictitious about these facts. Actually, casual observation as well as many legal and accounting features support these claims. In accordance with Freund (1897), unity, distinctiveness and identity through succession allow one to properly speak of collective capabilities, firm competitiveness, identity and reputation, and so on, as real properties of the firm as a real entity. Sets of contracts and collections of assets simply do not have any of these properties.[9]

It follows that the theory of the firm cannot exclude certain important "holistic" aspects. This does not imply any form of determinism or collectivism that annihilates individual agency, since wholes do not fully determine their parts and members any more than parts and members fully determine wholes. It simply means that certain structural, functional and systemic considerations need to enter the picture. Indeed, "a system . . . is . . . a complex thing whose components are bound together, as a consequence of which the whole has peculiar properties and behaves as a unit in some respects" (Bunge 2000: 148). Such "emergent properties" are properties of the whole not reducible to properties of its constituents. Given that emergence is by definition a bottom-up relation between ontological levels, it is fully relevant to the theory of the firm as a real entity.[10]

Importantly, emergence entails novelty. As a causal process, emergence accounts for new properties at the emergent level. Therefore, we need to understand the crucial and complex links between emergent entities and their properties and lower-level entities and their properties. In this respect, what does it mean to say that the higher-level entity is real? For one, this implies that it is persistent: "emergents appear as integrated wholes that tend to maintain some sense of identity over time" (Goldstein 1999: 50). For another, it implies that this higher level has some sort of causal power. Emergent levels exert some form of "downward causation" (see Emmeche *et al.*, 2000) or a form of "reconstitutive downward causation" (Hodgson 2003) that partly mould lower levels by giving shape and purpose to the interacting constituents. Again, this does not mean that higher levels fully determine lower levels but that this complex two-way causality of what may be termed "constitutive dynamics" is part of the "ontological glue" that keeps the whole together.

We may summarise our position with Dejnožka's (2006) reminder that existence is power for Plato: power to cause, power to stay unified, power to avoid dissolution. Fictitious or conceptual entities do not have such causal powers. For Leibniz, the stronger the unity or wholeness, the more real the entity. The

stronger the unity or the cohesiveness of the whole, the more we move from random sets to real entities (French 1982; Copp 1984; Tuomela 1989; Dejnožka 2006). In this sense, simple aggregates or random sums are to be distinguished from entities that deserve to be considered "real." One may thus properly distinguish "aggregates" from "social integrates" (e.g., Pettit 2003; Copp 2006). Aggregates simply do not have any of the properties of real entities. Given these ontological considerations, we reject the fiction theory that eliminates the reality of the firm as a whole and we reject the aggregate theory that reduces the whole to some of its members or parts.

### Cohesiveness and durability through time

From a theoretical point of view, firms as wholes have human members and non-human parts arranged into a complex structure or interactive system. By adding precision to the central concept of the "glue" that holds parts and members together as a cohesive whole, our discussion stresses that unity, wholeness and cohesiveness as well as identity and durability through time are essential ingredients of a theory of real entities. In his search for new foundations, Zingales (2000) feels that these issues are important but the ontology involved remains atomistic. An important point that reductionists miss is that once a certain cohesion and unity is attained, group behaviour is not identical or reducible to the behaviours of its elements. In fact, "groups may be cohesive, which individuals cannot be, and cohesiveness may affect the stability of the group, which is again something individuals … cannot have" (Brodbeck 1958: 16). Cohesion and unity can properly be said to be properties of the firm.

From our perspective, a theory of the firm as a real entity needs to include some notion of "ontological glue" understood here as being at least in part a form of efficient causation (in Aristotle's terms), i.e., as the complex system of upward causation and reconstitutive downward causation. Weissman (2000) appositely speaks of "causal reciprocity." We also need to include "teleological glue" as a form of final causation (again, in Aristotle's terms) manifested among other things by collective purpose and forward-looking behaviour. The firm's collective action capacity is reinforced by: "institutional glue" provided both by formal and informal rules, and by habits and routines; "organisational glue" manifested by structures, processes, functions and roles; "motivational glue" that ensures adherence to the common goal; "cognitive glue" accounting for shared beliefs and representations; and, finally, "productive synergy glue" or "competence glue" which relates to the complex co-specialisation of human assets such as knowledge and non-human assets. All these blends of glue contribute to the firm's cohesion and organise the collective action of its members, that is, the firm's relative behavioural unity through time. The more the emergent whole is unified or integrated, the more it can and should be considered a real entity.

Ontological individualism is based on the commonplace that collective entities such as firms cannot exist without any individuals. Yet the firm's persistence through time, based on the replication of behavioural patterns and collective

routines, implies that any particular firm continues to exist even if all its present human constituents are progressively replaced. Such independence qualifies the firm as a real entity: "economic entities ... actually have reality, in the sense that they are patterns which exist independently of their parts" (Raby 1959: 460). Our discussion fits a crucial empirical and theoretical insight from the evolutionary theory of the firm, namely that collective knowledge is typically "sticky" and that collective competence is retained through progressive change in firm membership. If this were not the case, the firm would certainly be a flimsy entity, incapable of survival in an evolutionary setting and thus incapable of preventing dissolution. Indeed, in terms of the "units of selection" debate in evolutionary economics, "group selection operates when the individuals in the group are bound together in a sufficiently cohesive manner to share a mostly common fate" (Hodgson and Knudsen 2004: 300).[11]

Ideas similar to those presented here can be found in the literature. For instance, Langlois and Foss (1997: 213) argue that firms are not held together by the "thin glue of transaction-cost minimization, but rather by the thicker glue of capabilities." Kay (2000: 704) considers that "[the] glue is likely to be found in complex strategic decisions made by the firm." Likewise, philosophers such as Pettit (2003) and Copp (2006) demonstrate that some groups typically "collectivise reason" in many complex decision-making situations and that this institutionalisation may act as a glue. We can thus properly speak of "integrates" as opposed to "aggregates." Finally, Commons (1924) and Raby (1959) both insist on what may be called a "working rules glue" combined with "expectational glue" characteristic of any "going concern":

> That which holds the going concern together is [a set] of working rules affording an expectation of a gross income to be obtained jointly ... If the expectation fails, the immortality fails. When the expectation continues, the corporation is a 'going concern.' For this reason, the legal form is subordinate.
>
> (Commons 1924: 145)

It is the real, ongoing collective dynamics of activity bound together by forward-looking behaviour patterns that is the basis of many of the real entity's properties and causal powers manifested by its economic process. Although legal form is subordinate to economic substance, as real entity theorists correctly stressed, it is important to understand the role played more generally by institutions, considering that the "circular, positive feedback from institutions to individuals and from individuals to institutions can help enhance the durability of the institutional unit" (Hodgson 2003: 172).

### *Institutions and individuals*

Institutions are both constraints and resources for individuals. Institutions both depend on individuals and channel individuals' behaviour. Institutions depend

on real acts, and are real because of these acts. In this respect, the problem of structure and agency (Swanson 1992; Archer 1995; Lawson 2003; Hodgson 2004) is an ontologically crucial one for social science in general and economic theory in particular. Our arguments in this chapter are in line with this literature's fundamental insight that social reality is composed of something more than, and is non-reducible to, individuals. Structure is a product of, a constraint on and a resource for individual and collective agency. One can properly speak of embodied institutional reality. Thus, in her ontology of institutional kinds, Thomasson (2003: 605) writes: "although they do not meet the [strong] realist paradigm of entities entirely independent of us for their existence and essence, they certainly are also not mere mental constructs." The firm as a real institution exists independently of its individual members but does not function apart from them.

In order to function, the firm relies on the structured dynamics of real acts of individuals and groups of individuals. In fact, internal organisation of authority and competence allows one to attribute "secondary actions" to collective entities such as firms via "primary actions" of their individual members in specific corporate positions (Copp 1979). In other words, corporations "act" through their internal decision structure (French 1982). From a different perspective, Searle's (1995) "constitutive rule" that "X counts as Y in context C" can be directly applied to individual and collective action: action X of corporate officer *x* counts as action Y of the firm because of *x*'s corporate position. Any consequent problem that may arise is therefore that of the firm as a whole as long as *x* acted in accordance with the "business judgment rule" or any other such rule. There are, therefore, ways of legally and conceptually attributing actions directly to the firm without "personifying" or "hypostatising" the firm. The firm "acts" not literally but institutionally through actors-in-the-mask. This is similar to Freund's (1897) representation principle that links individual agency to collective agency.

More fundamentally for our discussion, as a real economic entity, the firm is an organisation comprising human members and non-human parts that can be seen as an active "managed economic system" (MacMillan and Farmer 1979; Biondi 2005). Inside firms, humans typically engage in various activities the composite outcomes of which constitute what may properly be said to be the firm's particular productive activities. Intra-firm organisational and institutional structure is the complex combination of typically rule-based, goal-intended and forward-looking behaviour. Although collective routines ensure the replication of outcomes, the replication is imperfect, and this possibility, among others, introduces variation and heterogeneity in both organisational and productive outcomes. Accordingly, the need to maintain relative behavioural unity between different periods of time is essential to the firm's coherence. Hence, commitment to a common goal or undertaking is crucial in order for the relative behavioural unity to emerge and persist through time.

It is important to recognise that people do things together, and that people inside firms and other organisations reason in terms of "we" and "they" (Simon 2002). Such "identification is a powerful force ... by virtue of the loyalty it can produce to the goals of the whole system" (Simon 1991: 41, reprinted in this

volume). The firm's identity is an emergent cognitive property that certainly has causal powers of reconstitutive downward causation. The possibility of conflicts or power struggles and other incentive compatibility problems does not change our theoretical analysis of the nature of the firm as a real complex and active whole in significant ways. In fact, our analysis amplifies the importance of institutions in achieving the relative behavioural unity of collective action allowing the firm as a whole to attain its intended goals even though these may change or be betrayed.

When we say that what holds the firm together is a combination of a variety of types of glue, we do not mean to imply that all firms are identical in their actual unity. Building up this unity is a dynamic and temporal process and the stronger the cohesiveness or wholeness the more real the entity. Our view is thus compatible with Biondi's (2005) concept of the "becoming concern." Following Simon's (1996) discussion of complexity, artificiality and emergence, firms function and survive in their environment because they are examples of "adaptive artifice." Given this complexity, it is impossible to agree with Jensen and Meckling's (1976: 50) view that "the 'behavior' of the firm is like the behavior of the market," since the concept of "equilibrium" cannot apply to the firm as a whole (Biondi 2005). The firm is not some "spontaneous order" guided by an "invisible hand."

We follow Dejnožka's (2006) important ontological discussion in concluding that firms as real entities are both "interactive systems" and "complex artefacts." As interactive systems, firms function through the agential power of their members. As artefacts, firms typically exhibit purposive design and institutionalised functional unity. When we say that firms are artefacts, we do not imply that they can be reduced to instruments in the hands of shareholders or of any other constituents, or that their design is flawless. We simply mean that they are non-natural entities that are a product of human design and deliberate action. The members' agential power, or the deliberate determination of people to make their firm succeed, given certain goals and rules of action, is necessarily part of the ontological glue that accounts for the firm's cohesiveness and durability through time. Adopting a non-reductionist view on the nature of the firm does not eliminate individual agency. The contrary is quite true.

## Implications and applications

We have theoretically outlined what it means to say that the firm is not a fiction, legal or otherwise, but a real entity that typically exhibits a relative degree of cohesiveness and durability through time. The entity is real in that it possesses causal power to resist dissolution – not mechanically, but through the agential power of its ongoing human constituents, whose real actions are organised and to some extent institutionalised. The firm's ontological glue is causal, teleological, institutional, organisational, motivational, cognitive, expectational, and productive- or competence-based. The overall result is the coherence and continuity of the whole which can no longer be viewed as reducible to any of its constituent parts or members. It follows that reductionism and

more generally ontological individualism lead to a failure of the theory of the firm and of other social organisations. The nexus of contracts and collection of assets theories of the firm fail in their account of the firm on the grounds provided throughout this chapter. Our discussion carries several implications for current debates regarding the theory of the firm as well as some direct applications to group regulation.

### *Implications for the theory of the firm*

The theory of the firm is traditionally taken to be about three questions: the existence and nature of the firm (although these are two different questions); the boundaries of the firm; and the internal organisation of the firm. We believe that our account of the nature of the firm sheds new light on what should properly be considered a fourth question of a theory of the firm, namely the question of firm heterogeneity. Indeed, talk of the firm's distinctiveness, identity or individuality implies this fourth question. Our account of the cohesiveness and wholeness of the firm also relates to the second sense of the term "individuality": relative indivisibility or, put differently, relative non-decomposability. This does not mean that economic analysis should never take a closer look at particular constituents or relations. We are not denying that contracts are relevant for the theory of the firm. What it does mean, however, is that the whole should not be forgotten or explained systematically away.

Methodological individualism, as a monist doctrine of the "unique" or "ultimate" mode of analysis, fails if applied systematically. Many aspects of the firm such as competence and knowledge that rightly should be part of a theory of the firm are collective or holistic in nature. Ignoring them has contributed to the failure of the mainstream theory of the firm. In his discussion of the current debate between the contractual and the competence theories of the firm, Williamson (1999: 1102) writes: the "troublesome argument is that of aggregation. Taking a more holistic view, the firm as a whole is different from and larger than the sum of the parts . . . Inasmuch as transaction cost economics purports to be interested in all regularities whatsoever, it stands to benefit from research in the competence tradition on holistic consequences." Although we have not discussed the transaction cost economics theory of the firm as a "governance structure," Williamson's point applies in general to all contractual theories that hold that the contract or the transaction is the basic unit of analysis.

In line with a growing recent literature, we consider that combining the more holistic competence perspective with the contractual approach is one of the main challenges awaiting the theory of the firm (see Weinstein, this volume). Gibbons (2005: 238) has expressed this view under the heading "breaking loose." The contractual approach generally examines isolated transactions such as given "make-or-buy" situations, ignoring the link between this specific transaction and all other economic activities of the firm. On the other hand, the competence perspective insists on firms as non-reducible dynamic wholes. As such, the firm is a productive entity, a "competent team" (Eliasson 1990) and a

"knowledge-creating entity" (Nonaka, Toyama and Nagata 2000). It is as a whole that the firm is competitive or not. Both levels of individual incentives and contracts, and of collective knowledge, learning and competence, need to enter the picture. The question of combining the two perspectives on the theory of the firm is a fundamentally ontological undertaking that relies on the recognition that reality is multi-layered. By providing an account of the causal reciprocity between levels, our analysis of the cohesiveness and durability of the firm through time reinforces the relevance for the theory of the firm of the evolutionary ontology developed in the "units of selection" debate (e.g., Hodgson and Knudsen 2004).

In order to further explain the heterogeneity of firms, the approach needs to rely on both universal and historically specific concepts, and these need to be coherently articulated (Hodgson 1998). We have argued that unity, cohesiveness and durability through time of a complex combination of members and parts are the ontological features of the firm as a real entity. This ontological skeleton needs to receive economic, legal and accounting flesh, i.e., to be contextualised in terms of the particular economic, legal and accounting flesh attributes a specific firm has. The particular constellation of intra-firm and extra-firm relationships characterising each given firm depends on historical circumstances (Chandler 1992). We believe that our general theory of the firm as an economic entity is logically compatible with the observed varieties of capitalism, since it admits indefinitely many blends and varieties of teleological, institutional and organisational unity. The particular type of legal entity a given firm happens to be provides form only, the substance of the firm being independent of legal artifice.

### *Applications to integrated multi-unit structures*

Our general view according to which the stronger the unity or cohesiveness or wholeness, the more real the entity, has direct applications to integrated multi-unit structures such as business groups and vertical forms of network organisation, both from a theoretical point of view and from the point of view of regulation. Blumberg (1990b, 1993) has argued at length that "enterprise law" should apply to complex structures such as corporate groups, meaning that law should not systematically place form over substance or legal entity over real enterprise (see Berle 1947, reprinted in this volume, and Manfrin, this volume). Indeed, applied directly, corporate law tends to see multiple entities instead of one whole economic enterprise. However, in terms of liabilities against outsiders such as tort victims or final consumers, such multi-unit structures need to be treated as single economic wholes. For Deakin (2003), the very idea of "enterprise risk" needs to be reconsidered in the light of these complex structures. In fact, Deakin argues, a general move towards "organisational liability" within tort law is needed. The implicit ontological commitment underlying these views is very different from the fiction and aggregate theories discussed in this chapter, since for contractarians such as Easterbrook and Fischel (1985: 89), "the liability of 'the corporation' is limited by the fact that the

corporation is not real." By the same token, a corporate group is presumably not real from this point of view.

The implications of our real entity theory are clear in relation to this problem. Internal legal independence between subsidiaries should not blind us to the fact that these complex structures are complex artefacts designed to compete collectively with other such enterprises as wholes. The fact that each unit composing the structure may or may not be pursuing other objectives does not change the analysis. In the same spirit, Orts (1998) has argued that what he calls "relational firms," i.e., those complex structures that maintain internal legal autonomy while remaining more or less cohesive economic units, should legally be treated accordingly, depending on the question at hand. For Orts (1998: 313), "whether a relational firm is considered one entity or a group of entities will depend . . . on the question that is asked . . . If the question is whether a relational firm is acting as a unit in competing with other firms in markets, the answer will focus on the lines of authority that enable it to act as a single entity."

Blumberg's corporate groups are just one form of Orts's relational firms, one in which equity participation and controlling interest in subsidiaries are a major feature of the parts–whole relationship. Corporate groups, conglomerates and holding companies all share this feature in different degrees and for different purposes, varying from majority to minority ownership structures. The question of applying our arguments to multi-unit forms of economic organisation that do not involve equity ownership is even more challenging. We have, in fact, applied some of the arguments of this chapter to vertical forms of network organisation in which a "hub-firm" or "vertical architect firm" organises production between legally independent entities (Baudry and Gindis 2005). In accordance with Orts, we argued that the hub-firm's power to control the network as a whole and to prevent its dissolution implies that a vertical network form of this sort may properly be treated as a single economic entity.

In all these interactive complex systems, be they single firms, business groups or vertical network forms of organisation, the bottom line is that the stronger the unity and durability of action, the more real the economic entity, regardless of formal legal structure. The whole can be (and is) held together by the various forms of glue we have described, and it seems trivial to state that such wholes often thrive by competing with each other.

## Acknowledgment

I would like to thank Yuri Biondi, Jan Dejnožka and Geoffrey Hodgson for their valuable comments on previous versions of this essay. All remaining errors and shortcomings are my own.

## Notes

1 A point of view also adopted by Friedman (1970) and restated recently by Jensen (2002). Buchanan (1988: 136) thus reminds us that "the values or interests of individuals

are the only values that matter for the quite simple reason that these are the only values that exist."

2 Given that the contractual association is voluntary, the firm has nothing to do with state grants or statutes. Summing up this position, Butler and Ribstein (1989: 768) celebrate the fact that "market forces have long since freed the corporation from the control of state legislators, and the corporation can now safely be viewed as fundamentally contractual." Easterbrook and Fischel (1991) and countless other private ordering proponents also hold this anti-regulatory view. In this chapter, we ignore this normative issue and focus on the ontological commitments of the contractual theory only.

3 For detailed presentations of the corporate entity debate in the U.S. context, see Mark (1987), Hovenkamp (1988), Bratton (1989b), Hager (1989), Blumberg (1990a), Millon (1990), Horwitz (1992), Phillips (1994), Iwai (1999) and Avi-Yonah and Sivan (this volume). See Teubner (1988) for a general comparative approach and Foster (2000) for a comparison between England and France. For detailed historical accounts of business economics and entity theory debates in both Continental Europe and the New World, see Canziani and Kirat in this volume.

4 In this chapter, we ignore the "concession theory of the corporation" and the question of state involvement and focus on the "fiction vs. reality" and "aggregate vs. entity" issues only.

5 With the advent of LLCs, LLPs, LLLPs, and so forth, limited liability is no longer a unique characteristic of corporations, and traditional differences between partnerships and corporations have become less obvious. One-man limited liability companies exist in many jurisdictions. Some jurisdictions allow various "check-the-box" schemes for setting up a business where one may literally choose a menu of items from different statutes if their combination is the most suited for the venture. We are not interested in these evolutions *per se* but in the shift of emphasis they provoke in the debate about "entity" status, from corporations to all forms of business companies. In this context, some have expressed the need for a "unified business entity code" applying to all legal forms (e.g., Blackwell 1999). Private ordering proponents such as Ribstein (2003) are of course very critical of these calls for "entity rationalisation," since this could only be done through regulation.

6 Strikingly, however, debates often centre on the "sacred cow" of corporate existence (Israels 1952). Litowitz (2005: 501) writes that "from the perspective of cultural theory . . . the modern corporation is fundamentally a religious and mythological entity." We have before us a battle of the gods between "the corporation" and "the shareholders." A serious effort of "demystification" of the corporate entity is needed.

7 See, in this volume, Moore and Rebérioux on entity theory and corporate governance, and Avi-Yonah and Sivan on entity theory and corporate social responsibility.

8 Entity theories were themselves much criticised at the time. For instance, Singleton (1912: 291) considers that entity views are "ontological theories incapable of verification," Wormser (1912: 496) avoids the "tempting but profitless discussion," Radin (1932: 658) brushes off the debate as a "matter of literature" and Cohen (1935) speaks mockingly of "transcendental nonsense." The ontological individualists discussed in the second section are saying roughly the same thing as are aggregate theorists such as Morawetz (1882).

9 Hart (2001: 1714) recognises that "somehow there has to be some stickiness in the firm or system, so that a firm's reputation can be separated from that of key personnel." Hart's intuition is again correct, although he does not provide an adequate theory of this stickiness.

10 This is quite different from the widely accepted thesis of "supervenience" which contains nothing in it to make it a necessary vertical relation (Humphreys 1997). Even if applied to a vertical relation between levels, supervenience retains the ontological priority of the lower level: once the facts about the lower-level entities are set, so too are all the aggregate facts (Kincaid 1995). There is still something missing from an

ontological point of view. On the other hand, emergence is a failure of aggregativity (Wimsatt 1997) and contains an explicit ontological commitment.

11 Legally, firm members are certainly bound together, since they share the common fate of the firm in the case of dissolution or absorption. Human asset specificity and knowledge and productive synergies are also by definition dependent on the firm's survival, since their value is greater or only exists inside the firm.

## References


Alchian, Armen A. (1984), "Specificity, Specialization and Coalitions," *Journal of Institutional and Theoretical Economics*, 140(1): 34–49.

Alchian, Armen A. and Demsetz, Harold (1972), "Production, Information Costs and Economic Organization," *American Economic Review*, 62(5): 777–95.

Alchian, Armen A. and Woodward, Susan (1988), "The Firm Is Dead; Long Live The Firm: A Review of Oliver E. Williamson's The Economic Institutions of Capitalism," *Journal of Economic Literature*, 26(1): 65–79.

Allen, William T. (1992), "Our Schizophrenic Conception of the Business Corporation," *Cardozo Law Review*, 14(2): 261–81.

Archer, Margaret S. (1995), *Realist Social Theory: The Morphogenetic Approach*, Cambridge: Cambridge University Press.

Baudry, Bernard and Gindis, David (2005), "The V-Network Form: Economic Organization and the Theory of the Firm," unpublished, U. Lyon 2. Available online at http://ssrn.com/abstract=795244 (accessed 16 September 2006).

Berle, Adolf A. (1947), "The Theory of Enterprise Entity," *Columbia Law Review*, 47(3): 343–58, reprinted in this volume.

Biondi, Yuri (2005), "The Firm as an Entity: Management, Organisation, Accounting," paper no. 46, U. Brescia. Available online at http://ssrn.com/abstract=774764 (accessed 16 September 2006).

Blackwell, Thomas F. (1999), "The Revolution Is Here: The Promise of a Unified Business Entity Code," *Journal of Corporation Law*, 24(2): 333–78.

Blair, Margaret M. (1995), *Ownership and Control: Rethinking Corporate Governance for the Twenty-First Century*, Washington D.C.: Brookings Institution Press.

—— (1999), "Firm-Specific Human Capital and Theories of the Firm," in Margaret M. Blair and Marc J. Roe (eds), *Employees and Corporate Governance*, Washington D.C.: Brookings Institution Press.

Blumberg, Phillip I. (1990a), "The Corporate Personality in American Law: A Summary Review," *American Journal of Comparative Law*, 38(S): 49–69.

—— (1990b), "The Corporate Entity in an Era of Multinational Corporations," *Delaware Journal of Corporate Law*, 15(2): 283–375.

—— (1993), *The Multinational Challenge to Corporation Law: The Search for aNew Corporate Personality*, New York: Oxford University Press.

Bratton, William W. (1989a), "The 'Nexus-of-Contracts' Corporation: A Critical Appraisal," *Cornell Law Review*, 74(2): 407–65.

—— (1989b), "The New Economic Theory of the Firm: Critical Perspectives from History," *Stanford Law Review*, 41(6): 1471–527.

Brodbeck, May (1958), "Methodological Individualisms: Definition and Reduction," *Philosophy of Science*, 25(1): 1–22.

Buchanan, James M. (1988), "Contractarian Political Economy and Constitutional Interpretation," *American Economic Review*, 78(2): 135–9.

Bunge, Mario (2000), "Ten Modes of Individualism – None of Which Works – And Their Alternatives," *Philosophy of the Social Sciences*, 30(3): 384–406.

—— (2003), *Emergence and Convergence: Qualitative Novelty and the Unity of Knowledge*, Toronto: University of Toronto Press.

Butler, Henry N. and Ribstein, Larry E. (1989), "The Contract Clause and the Corporation," *Brooklyn Law Review*, 55(4): 767–808.

Chandler, Alfred D. (1992), "What Is a Firm? A Historical Perspective," *European Journal of Economics*, 36(2–3): 483–92.

Cheung, Steven N.S. (1983), "The Contractual Nature of the Firm," *Journal of Law and Economics*, 26(1): 1–21.

Chisholm, Roderick M. (1973), "Parts as Essential to Their Wholes," *Review of Metaphysics*, 26(4): 581–603.

Coase, Ronald H. (1937), "The Nature of the Firm," *Economica* n.s., 4(16): 386–405.

Cohen, Felix S. (1935), "Transcendental Nonsense and the Functional Approach," *Columbia Law Review*, 35(6): 809–49.

Commons, John R. (1924), *Legal Foundations of Capitalism*, New York: Macmillan. Reprinted 1995, New Brunswick, NJ: Transaction Publishers.

Copeland, Morris A. (1927), "An Instrumental View of the Part–Whole Relation," *Journal of Philosophy*, 24(4): 96–104.

Copp, David (1979), "Collective Actions and Secondary Actions," *American Philosophical Quarterly*, 16(3): 177–86.

—— (1984), "What Collectives Are: Agency, Individualism and Legal Theory," *Dialogue*, 23(2): 249–69.

—— (2006), "On the Agency of Certain Collective Entities: An Argument from 'Normative Autonomy'," *Midwest Studies in Philosophy*, 30: 184–221.

Deakin, Simon (2003), "'Enterprise-Risk': The Juridical Nature of the Firm Revisited," *Industrial Law Journal*, 32(2): 97–113.

Dejnožka, Jan (2006), *Corporate Entity*, unpublished manuscript. Available online at http://www.members.tripod.com/~Jan_Dejnozka/corporate_entity_book.pdf (accessed 15 September 2006)

Demsetz, Harold (1988), "The Theory of the Firm Revisited," *Journal of Law, Economics and Organization*, 4(1): 141–61.

Dodd, E. Merrick (1932), "For Whom Are Corporate Managers Trustees?" *Harvard Law Review*, 45(7): 1145–63.

Donaldson, Thomas and Preston, Lee E. (1995), "The Stakeholder Theory of the Corporation: Concepts, Evidence and Implications," *Academy of Management Review*, 20(1): 65–91.

Easterbrook, Frank H. and Fischel, Daniel R. (1985), "Limited Liability and the Corporation," *University of Chicago Law Review*, 52(1): 89–117.

—— (1991), *The Economic Structure of Corporate Law*, Cambridge, MA: Harvard University Press

Eisenberg, Melvin A. (1999), "The Conception That the Corporation is a Nexus of Contracts and the Dual Nature of the Firm," *Journal of Corporation Law*, 24(4): 819–36.

Eliasson, Gunnar (1990), "The Firm as a Competent Team," *Journal of Economic Behavior and Organization*, 13(3): 275–98.

Emmeche, Claus, Køppe, Simon and Stjernfelt, Frederik (2000), "Levels, Emergence and Three Versions of Downward Causation," in Peter B. Anderson *et al.* (eds), *Downward Causation: Minds, Bodies and Matter*, Arhus: Arhus University Press.

Fama, Eugene F. (1980), "Agency Problems and the Theory of the Firm," *Journal of Political Economy*, 88(2): 288–307.

Foster, Nicholas H.D. (2000), "Company Law Theory in Comparative Perspective: England and France," *American Journal of Comparative Law*, 48(4): 573–621.

French, Peter A. (1982), "Crowds and Corporations," *American Philosophical Quarterly*, 19(3): 271–7.

Freund, Ernst (1897), *The Legal Nature of Corporations*, Chicago: Chicago University Press.

Friedman, Milton (1970), "The Social Responsibility of Business Is to Increase Its Profits," *New York Times Magazine*, September 13.

Gibbons, Robert (2005), "Four Formal(izable) Theories of the Firm?," *Journal of Economic Behavior and Organization*, 58(2): 200–45.

Goldstein, Jeffrey (1999), "Emergence as a Construct: History and Issues," *Emergence*, 1(1): 49–72.

Greenwood, Daniel J.H. (1996), "Fictional Shareholders: For Whom Are Corporate Managers Trustees Revisited," *Southern California Law Review*, 69(5): 1021–104.

—— (2005), "Introduction to the Metaphors of Corporate Law," *Seattle Journal for Social Justice*, 4(1): 1–31.

Grossman, Sanford and Hart, Oliver D. (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94(2): 691–719.

Hager, Mark M. (1989), "Bodies Politic: The Progressive History of Organizational 'Real Entity' Theory," *University of Pittsburgh Law Review*, 50(3): 575–654.

Hansmann, Henry (1996), *The Ownership of Enterprise*, Cambridge, MA: Harvard University Press.

Hart, Oliver D. (1989), "An Economist's Perspective on the Theory of the Firm," *Columbia Law Review*, 89(7): 1757–74.

—— (1995), *Firms, Contracts and Financial Structure*, New York: Oxford University Press.

—— (2001), "Norms and the Theory of the Firm," *University of Pennsylvania Law Review*, 149(6): 1701–15.

Hart, Oliver D. and Moore, John (1990), "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98(6): 1119–58.

Hodgson, Geoffrey M. (1998), "The Coasean Tangle: The Nature of the Firm and the Problem of Historical Specificity," in Steven G. Medema (ed.), *Coasean Economics: Law and Economics and the New Institutional Economics*, Boston: Kluwer.

—— (2002), "The Legal Nature of the Firm and the Myth of the Firm-Market Hybrid," *International Journal of the Economics of Business*, 9(1): 36–60.

—— (2003), "The Hidden Persuaders: Institutions and Individuals in Economic Theory," *Cambridge Journal of Economics*, 27(2): 159–75.

—— (2004), *The Evolution of Institutional Economics: Agency, Structure and Darwinism in American Institutionalism*, London: Routledge.

Hodgson, Geoffrey M. and Knudsen, Thorbjørn (2004), "The Firm as an Interactor: Firms as Vehicles for Habits and Routines," *Journal of Evolutionary Economics*, 14(2): 281–307.

Horwitz, Morton J. (1992), *The Transformation of American Law, 1870–1960: The Crisis of Legal Orthodoxy*, New York: Oxford University Press.

Hovenkamp, Herbert (1988), "The Classical Corporation in American Legal Thought," *Georgetown Law Journal*, 76(5): 1593–689.

Humphreys, Paul (1997), "Emergence, Not Supervenience," *Philosophy of Science*, 64(S): S337–S345.

Ireland, Paddy (1999), "Company Law and the Myth of Shareholder Ownership," *Modern Law Review*, 52(1): 32–57.

Israels, Carlos L. (1952), "The Sacred Cow of Corporate Existence: Problems of Deadlock and Dissolution," *University of Chicago Law Review*, 19(4): 778–93.

Iwai, Katsuhito (1999), "Persons, Things and Corporations: The Corporate Personality Controversy and Comparative Corporate Governance," *American Journal of Comparative Law*, 47(4): 583–632.

Jensen, Michael C. (2002), "Value Maximization, Stakeholder Theory and the Corporate Objective Function," *Business Ethics Quarterly*, 12(2): 235–56.

Jensen, Michael C. and Meckling, William (1976), "Theory of the Firm: Managerial Behavior, Agency Costs and Capital Structure," *Journal of Financial Economics*, 3(4): 305–60.

Kay, Neil M. (2000), "Searching for the Firm: The Role of Decision in the Economics of Organization," *Industrial and Corporate Change*, 9(4): 683–707.

Khalil, Elias L. (1996), "Is the Firm an Individual?," *Cambridge Journal of Economics*, 21(4): 519–44.

Kincaid, Harold (1995), "The Empirical Presuppositions of Metaphysical Explanations in Economics," *Monist*, 78(3): 368–85. Reprinted in Uskali Mäki (2001) (ed.), *The Economic World View: Studies in the Ontology of Economics*, Cambridge: Cambridge University Press.

Langlois, Richard N. and Foss, Nicolai J. (1997), "Capabilities and Governance: The Rebirth of Production in the Theory of Economic Organization," *Kyklos*, 52(2): 201–18.

Laski, Harold J. (1916), "The Personality of Associations," *Harvard Law Review*, 29(4): 404–26.

Lawson, Tony (2003), *Reorienting Economics*, London: Routledge.

Litowitz, Douglas (2005), "The Corporation as God," *Journal of Corporation Law*, 30(3): 501–38.

Machen, Arthur W. (1911), "Corporate Personality," *Harvard Law Review*, 24(4): 253–67 and 24(5): 347–65.

MacMillan, Keith and Farmer, David (1979), "Redefining the Boundaries of the Firm," *Journal of Industrial Economics*, 27(3): 277–85.

Maitland, Frederick W. (1900), "The Corporation Sole," *Law Quarterly Review*, 16(3): 335–54.

—— (1903), "The Unincorporated Body," in H.L.A. Fischer (1911) (ed.), *The Collected Papers of Frederick William Maitland*, vol. 3, Cambridge: Cambridge University Press.

—— (1905), "Moral Personality and Legal Personality," *Journal of the Society of Comparative Legislation*, 6(2), 192–200.

Mäki, Uskali (2001), "Economic Ontology: What? Why? How?" in Uskali Mäki (ed.), *The Economic World View: Studies in the Ontology of Economics*, Cambridge: Cambridge University Press.

Mark, Gregory A. (1987), "The Personification of the Business Corporation in American Law," *University of Chicago Law Review*, 54(4): 1441–83.

Millon, David (1990), "Theories of the Corporation," *Duke Law Journal*, 1990(2): 201–62.

Moore, John (1992), "The Firm as a Collection of Assets," *European Economic Review*, 36(2–3): 493–507.

Morawetz, Victor (1882), *A Treatise on the Law of Private Corporations Other Than Charitable*, Boston: Little, Brown.

Nonaka, Ikujiro, Toyama, Ryoko and Nagata, Akiya (2000), "A Firm as a Knowledge-Creating Entity: A New Perspective on the Theory of the Firm," *Industrial and Corporate Change*, 9(1): 1–20.

Orts, Eric W. (1998), "Shirking and Sharking: A Legal Theory of the Firm," *Yale Law and Policy Review*, 16(2): 265–329.

Pettit, Philip (2003), "Groups with Minds of their Own," in Frederick Schmitt (ed.), *Socializing Metaphysics: The Nature of Social Reality*, Lanham, MD: Rowman & Littlefield.

Phillips, Michael J. (1992), "Corporate Moral Personhood and Three Conceptions of the Corporation," *Business Ethics Quarterly*, 2(4): 435–59.

—— (1994), "Reappraising the Real Entity Theory of the Corporation," *Florida State University Law Review*, 21(4): 1061–122.

Raby, William L. (1959), "The Two Faces of Accounting," *Accounting Review*, 34(3): 452–61.

Radin, Max (1932), "The Endless Problem of Corporate Personality," *Columbia Law Review*, 32(4): 643–67.

Ribstein, Larry E. (2003), "Making Sense of Entity Rationalization," *Business Lawyer*, 58(3): 1023–42.

Rowley, Scott (1931), "The Individuality of Business Associations," *Virginia Law Review*, 17(6): 557–69.

Ruben, David-Hillel (1983), "Social Wholes and Parts," *Mind*, 92(366): 219–38.

Searle, John R. (1995), *The Construction of Social Reality*, New York: Free Press.

Simon, Herbert A. (1991), "Organizations and Markets," *Journal of Economic Perspectives*, 5(2): 25–44, reprinted in this volume.

—— (1996), *The Sciences of the Artificial*, 3rd edn, Cambridge, MA: MIT Press.

—— (2002), "We and They: The Human Urge to Identify with Groups," *Industrial and Corporate Change*, 11(3): 607–10.

Singleton, W.E. (1912), "Entities and Real and Artificial Persons," *Journal of the Society of Comparative Legislation*, 12(2): 291–8.

Smith, Thomas A. (2001), "The Use and Abuse of Corporate Personality," *Stanford Agora*, 2(1): 69–72.

Stauss, James H. (1944), "The Entrepreneur: The Firm," *Journal of Political Economy*, 52(2): 112–27, reprinted in this volume.

Stout, Lynn A. (2005), "On the Nature of Corporations," *University of Illinois Law Review*, 2005(1): 253–67.

Swanson, Guy E. (1992), "Doing Things Together: Some Basic Forms of Agency and Structure in Collective Action and Some Explanations," *Social Psychology Quarterly*, 55(2): 94–117.

Teubner, Gunther (1988), "Enterprise Corporatism: New Industrial Policy and the 'Essence' of the Legal Person," *American Journal of Comparative Law*, 36(1): 130–55.

Thomasson, Amie L. (2003), "Realism and Human Kinds," *Philosophy and Phenomenological Research*, 67(3): 580–609.

Tuomela, Raimo (1989), "Actions by Collectives," *Philosophical Perspectives*, 3: 471–96.

Vining, Joseph (1978), *Legal Identity: The Coming of Age of Public Law*, New Haven, CT: Yale University Press.

Weissman, David (2000), *A Social Ontology*, New Haven, CT: Yale University Press.

Williamson, Oliver E. (1991), "Comparative Economic Organization: The Analysis of Discrete Structural Forms," *Administrative Science Quarterly*, 36(2): 269–96.

—— (1999), "Strategy Research: Governance and Competence Perspectives," *Strategic Management Journal*, 20(12): 1098–108.

Wimsatt, William C. (1997), "Aggregativity: Reductive Heuristics for Finding Emergence," *Philosophy of Science*, 54(S): S372–S384.

Wormser, I. Maurice (1912), "Piercing the Veil of Corporate Entity," *Columbia Law Review*, 12(6): 496–518.

Zingales, Luigi (2000), "In Search of New Foundations," *Journal of Finance*, 55(4): 1623–53.

# 16 The firm as an entity and the law

## The economic substance, the legal forms

*Federico Manfrin*

### The economic substance of the firm within and beyond its legal forms

#### *The economic activity in search of legal forms*

From ancient times onward, firms – whether small limited partnerships or the medium-sized glass and ceramic factories of the Greek and Roman Mediterranean – have carried out the economic activity. Since then, they have however been unable to avoid being concerned by legal regulations, either commercial or labor laws, or others.

The scope of regulation increased in time with the growing complexity of society and its legal systems, from European antiquity to the Middle Ages to – finally – the turning point of the two East India Companies in the first half of the seventeenth century, when the limited liability company was born. After this, and all the more so after the radical innovations given by the Napoleonic Code, the economic activity of the firm was more and more governed from the legal point of view by the regulation of the legal forms that "enterprises" assume in their economic activities, such as proprietorships, partnerships, limited liability companies, corporations – all identified under the name "company" in this chapter.

In fact, however strict, registration is necessary in order to take on formal existence, but the firm as an economic activity pre-exists laws as well as comprehensively including them. From a historical point of view at least (down to the third quarter of the twentieth century), the firm was not actually regulated *per se* but subject to commercial law in general and to special laws and regulations (labor, finance, international exchanges, environmental) in addition. To be clearer, laws regulate the enterprise in its wide and changeable economic activity. But this recognition by law merely gives legal forms to an economic substance that is far from being completely synthesized by it. Any operation inside the firm finds its full meaning solely in the management system operating the enterprise and not in the legal forms of the state juridical system which often lag behind the development of the underlying economic substance.[1] The economic activity of the firm consists therefore in *libertas sub lege* (freedom under law) but

also *sine lege* (without law) or *contra legem* (against the law). *Sine lege*, for instance in the case of business groups, considering that companies which represent their subsidiaries are a legal product of the sixteenth century, while the group is the product of the entrepreneurial inventiveness of the twentieth century and is still not fully regulated by the majority of state legal systems. *Contra legem*, when the juridical forms – especially in the event that they establish limited liability – are merely an instrument to fulfill the economic interest of a coalition of minority but controlling constituencies – generally, managers and dominant shareholders – as financial bankruptcies testify (see Moore and Rebérioux, and Berle, this volume). To summarize, the state juridical system can not fully cover the whole economic system. Legal forms can shape just a part but not the whole enterprise: they can be a synecdoche for the economic reality of the further-reaching and broader enterprise that overlies the legal forms involved.

From this viewpoint, the governing role is led by managerial discretion and by the regulating role of the state legal and judicial systems which coexist within the legal-economic system generated by the firm. This requires a clear definition of the meaning given to the two terms of the relationship. Such a legal-economic connection has been identified with partial juridical categories by the two leading approaches which have created a synthesis between economics and law, which are the (a) economic analysis of the law, and (b) the law of the economy (Canziani 2001). These theories try to explain both enterprises and companies, but often confuse the names: sometimes due to erroneous understanding, sometimes as a consequence of misleading semantic interpretations by readers. In this context, enterprises and companies recall the concepts of substance and form so beloved of the accounting doctrine: the firm is a dynamic coordination of collective functions which can imply a given juridical structure of which company, group and other legal arrangements are the peculiar elements. In the same way, this comparison can be made taking into consideration the two following perspectives: *a parte objecti* (mode of ongoing existence) and *a parte subjecti* (goals). The "enterprise," *a parte objecti*, has to pursue economic sustainability and, *a parte subjecti*, has to fulfill the common interest of all its constituencies. A troubling difference exists here with the "company" in that, *a parte subjecti*, formally appears to purport the fulfillment of the particular interests of its associates (shareholders), and on the contrary, *a parte objecti*, is ultimately framed by the laws of the state legal system, given that these laws are intended to guarantee collective and public interests as well.

This chapter aims on the one hand to treat the connection between the law and the independent, historical and only partially regulated system of the economic entity. On the other hand, its objective is to criticize some of the usual approaches to the substance of the firm, approaches of a juridical nature and of partial epistemic value. In addition it aims at developing a more encompassing legal-economic idea of the firm: the everlasting interpretation of the firm as an independent going concern of holistic nature which intends to produce new incomes by way of business coordinations. Last, this idea will be applied to two special problems, namely (i) industrial groups, and (ii) common interest.

### *The registered firm as a semi-autonomous juridical system*

The firm-entity, as described in this book, according to modern legal systems is neither fully recognized from a legal point of view, nor can it be a subject of law in its existence as an economic institute.[2] Nevertheless, a certain evolution must be acknowledged: a new legal recognition of the firm has been developed over the last years by some special fields of the law such as corporate group regulation, labor law, and tax law. Recent developments are more and more interested in the whole structure of an economic organization, that is, in the enterprise beyond the legal boundaries and classifications of its legal forms.

Notwithstanding such considerations, the firm always assumes a juridical quality, since the interrelations that constitute its core are legally relevant and often adopt various and different juridical forms: from the social contract, to the labor contract, to public regulation (environmental, financial, industrial, and so on) and, when coordinated by a web of legal arrangements or in a group, from franchising contracts to distribution agreements. The interaction between the economic substance and such different juridical forms constitutes what we can define as the "legal-economic system" of the firm.

As Avi-Yonah and Sivan (this volume) point out, we are used to reading in law books[3] that companies are legal persons born through an act of the public authority[4] (artificial entity view), that they can sometimes be legal fictions to support economic individual actions (aggregate entity view), and, in some cases, that they are real (natural) entities involving legal processes and implications (real entity view). Immediately led by the first idea, we ultimately forget "artificial" and we begin to think of companies only as "persons."[5]

The time has come to put an end to this usual but misleading personification of the firm. Many theories try to find a definition of the firm, among the others Mises' individual action theory, Schutz's phenomenology, and Kirzner's theory of entrepreneurial discovery, but it seems that they have all read one page or another of the same book. In fact, a firm is a matrix, a net, a systemic coordination of functions and structures. And all these structures make it clear that the firm is not an artificial but an *objective reality*, a coordinating institution, i.e. it is an "entity," according to the semantic proposed by the continental tradition of early advanced business economics (Canziani, this volume).

As the theory of the enterprise entity developed by Berle (1947, reprinted in this volume) early recognized,

> the divergence between corporate theory and the underlying economic facts has occasioned a variety of problems (dealt with ad hoc by the courts) in which the theory of "artificial personality" simply did not work, and was consequently extended, disregarded, sometimes buttressed by further fiction, at others manipulated to get a convenient result.

For instance, this interpretation of the firm as an entity is the one which led courts to develop the institutes of *de facto* company and of corporate group liability.[6]

Commons (1924, p. 145) underlined that:

> the legal form is subordinate. The concern may exist as a partnership, a union, an association, a corporation, a cooperative (or a composite legal form as a group). The essential thing is the visible, tangible, going concern of persons, with its invisible, intangible behavior of the immediate and remote future stabilized by working rules.

Whenever a social body such a firm exists, whatever its complexity, regulation takes place within its boundaries, involving a whole order of powers and sanctions. For such a reason, any power which is effectively social and which is therefore organized transforms itself into an institution, i.e. it has recourse to institutional devices that "prompt, frame and enhance" the organized and instituted activity.[7]

Firms are therefore, at the same time, both organizations and institutions. They are modes of social organization in which individuals and resources coordinate themselves through internal norms and general mandatory rules. The entity is an institution or a social body, meaning that it is the manifestation of the social nature – not purely individual – of humankind.

As other institutions, the entity firm establishes the synthesis which frames the single actor, in which not only its activities, but its own position is regulated, together with a system of freedoms, rules, guarantees, and controls, giving structure to and unifying a series of elements which are discretionally individual ones.[8] Such institutional framing is therefore the first, original and essential expression of the law. The law expresses itself in institutions, and institutions exist as such because in some instances they encompass the law.

A firm undertakes its operations through external interactions which are socially relevant and involve public interest. They are thus regulated by legal dispositions (rules) – *ubi societas ibi ius* (where there is society, there is law). At the same time, a firm also operates by means of internal relations between the "working constituents" discussed by Biondi in this volume. The entity-firm and the working constituents come into existence *uno actu* (strictly together) and mutually interact. They create working norms that transform the firm into an institution from which the working norms themselves take their strength. In this way, the entity firm itself becomes a "legal-economic system," and the working norms, as the direct product of the functioning of this inner institutional system, regulate – with limitations – the relations which constitute the firm as an entity. As a consequence, private law issued by outer legal dispositions interacts with the internal legal-economic system and comes to constitute the complementary institutional part of those relations as they actually happen.[9] Whether proper to common or to civil law, legal rules are involved and modified by such an involvement in the legal-economic system generated by the entity firm, in the same way as the inside working norms are shaped by legal regulation. Finally, the internal working norms, together with the external legal dispositions led by the state juridical system, are neither superimposed on nor a property of the firm

itself, but they constitute its institutional structure – *ubi ius ibi societas* (where there is law, there is society). Any legal system is an institution, and inversely every institution is a legal system.

The relationship between the two concepts can be considered – contrary to Kelsen – absolute and transitive, i.e. an identity (Romano 1977). The "firm" as a legal-economic entity does not emerge by a procedure established by the law, it does not exist by reason of a pre-existent legal disposition: its emergence is a fact. This is different for the legal forms the entity involves, since, as mere forms, they find their own place in the law of the state, i.e. in the legal system the firm lives in.

This perspective raises the following question: what lies beyond the choice of the entity's legal forms? Four approaches are possible, i.e. the firm seen either as (a) a mere artificial person, (b) a nexus of contracts, (c) the object of property rights, or, last but not least, as (d) an ongoing economic coordination. The latter may be considered as a synthesis and a further advance on the others, and it represents a whole new perspective of the economic and legal implications of the firm as an economic entity. These four approaches are dealt with below in the second section.

## The firm as an entity versus its partial juridical approaches

### *The firm as a mere artificial person*

According to this doctrine the company exists because the law creates it, and within the limits of such a creation. The state, with the aim of regulating the firm's coordinations, establishes the juridical institution of the artificial personality which in turn, in order to achieve existence, requires administrative concession. This view underlines the fact that, under both civil and common law, a company must be registered with the relevant governmental body (chamber of commerce, register of companies, and so on).

However, this view seems artificial and misleading as a comprehensive perspective of the firm, since firms exist independently from the administrative authorization: they exist before such authorization and sometimes they continue to exist after their winding-up (for instance when, after legal termination, the economic entity still continues to produce, workers continue to perform their jobs, products continue to be sold to customers, in other words the economic coordinations continue to operate). It is interesting to see how far from this reality Chief Justice Marshall's words sound:

> A corporation is an artificial being, invisible, intangible, and existing only in contemplation of law. Being the mere creature of law, it possesses only those properties which the charter of its creation confers upon it, either expressly or as incidental to its very existence.[10]

Keeping a realistic viewpoint, on the contrary, courts progressively (i) applied a rule similar to the *substance over form* in accounting and attributed legal

responsibility to entities which *stricto iure* (according to the law) were not responsible, or (ii) erased the quality of moral person with limited responsibility in those cases where moral persons – even if they had received formal investiture by the competent public authorities – were utilized, for example, to commit offences, to hide fraudulent behavior behind company's veil, to infringe anti-trust dispositions or to operate within the gray area of hidden oligopolies.[11]

The view of the firm as an artificial person can also refer to the business operations piloted by free owners. In this case, it reflects an idea close to economic and political *laissez-faire*. In particular, according to this theory, companies and all legal institutes with a juridical personality are essentially an instrument for the consensual agreement of more people. Firms become a means for the conduct of essentially private affairs, they – the entities in general – are not real institutions, but express the shareholders' freedom. The artificial person becomes just a specific juridical type in the hands of the shareholders for organizing the activity.

In fact, this approach aims, more or less explicitly, at pinpointing the relevant relationship between the system of operations conducted under individual economic freedom and the power of regulation and control advocated by the state in the public interest.

According to the institutional perspective described here, organizations always exist on a social level, with the legal system limiting its own functions to their recognition. The associative system is an expression of individual freedom, finding representation by means of the related body: the will of members in an association is different from the will of the single subject, due to the fact they are acting as components of a whole. The juridical person, in this context, operates as a collective recognition of a collective expression of individual freedom – the *collective mask* emphasized by Gindis (this volume). It therefore represents a due act in a free economic system. Persons and resources obtain reciprocal coordination: this makes the organization a special reality in which each member does not accomplish self-standing actions, but involves himself in the whole without any overlapping. The internal inter-personal relations often become organizational relations, and the functioning rules express the operating of the entity.

In this context, the elements and structures of the firm – once bound together and coordinated – originate an institution. This means a "legal-economic system" having its own norms and its managerial constituent acting at the same time as norm-setter and as an "inner arbitrating settler" (Biondi 2006), i.e. with a whole coordination of formal structures, functions, procedures, tasks (and rites if necessary). We could say, as a partial conclusion, that the nature of "institution" we ascribed to the firm results from its internal legal-economic system, not from the way in which it is recognized – generally incompletely – by the laws of the state.[12] In this context, we can therefore affirm that the firm is an institution both partly originating in the managerial working norms, and partly deriving from state regulation.

### *The firm as a nexus of contracts*

The legal-economic approach based on the contractual nature of organizations is dominant at present. According to this doctrine (Coase 1937; Alchian and Demsetz 1972; Williamson 1985) an organization may be considered as a nexus of contracts between the organization and other (participating) parties.

From this perspective, the deed of incorporation is not an act that creates an entity (the artificial person view), but a document in which different covenants convene. The granting of the contractual power to the organization allows for the definition of a nexus of bilateral contracts (between the company and the suppliers, clients, employees, lenders, and so on) instead of a confusing web of dissimilar multilateral agreements among all the interested parties.

The implications of such an immaterial contractual theory of the organizations, based on the significance of the contractual nature of participation, are not irrelevant. Two points can be emphasized. First, the voluntary nature of the participation by individual subjects in the organization: the association with an organization by a party (lender, supplier, employee, and so on), in a context where the rationality of individuals is taken for granted, can be only the result of a cost-benefit analysis finalized to the maximization of an individual goal-function. Second, the possibility of reorganizing, changing or leaving the organization, simply by redefining its individual contracts.[13]

However, this reconstruction of the nature of formal organizations is open to dispute: the contractual approach, entirely focused on the nature of the formal organization as the reference for relevant transactions, neglects a number of features which nonetheless represent relevant constitutive parts. We will consider here that the “nexus of contracts view” can neither (i) express the economic and financial dynamics of the firm, nor (ii) represent the nature and role of the economic production of the firm as well as the institutional consequences of its legal-economic system, nor (iii) take into account the development of the firm from the point of view of both its economic entity and its institutional role.

The firm cannot be seen as a mere nexus of contracts, since the law is not able to explain the very nature of the firm’s assets, as Zappa (1937: 111), quoting Beigel (1900: 52), underlined:

> The instant of the entry takes place, not when an agreement is juridically effective, but in the moment an assignment has been carried out, since commercial moments and not juridical transactions are the object of the record. But not even the moment of the transfer of a commercial asset possessed by a third party (in juridical terms the *traditio*) is relevant for the recording of the items, since only the change of possession which has taken place for its own firm is due to be considered, . . . The execution and not the agreement (or the stipulation) is always recorded. Any entry, drawn up by whatever methods, can not follow other principles.

The law alone is not capable of explaining the structure of accounts and, more generally, of interpreting any economic coordination *sub specie juris* (strictly

according to the law) (point (i)). For instance, an asset stolen but used by the firm belongs to the entity process, even though no contract can justify its possession. Moreover, accounting suggests substance over form by referring to actual economic control and implications instead of legal ownership in order to define assets belonging to an entity (Biondi, this volume).

With regard to the understanding of production and institutions (point (ii)), we find a theoretical progression from the neoclassical school, which can be identified with the model-structure of principal(s)-agent(s),[14] to the neo-institutional school, which in turn can be represented by Williamson (1975; 1985; 1991). According to the neoclassical approach, the problem of coordination is defined and solved by the "optimal configuration" of incentives: the firm becomes a set of multilateral agreements set up by a blind equilibrium machinery so as to reconcile individual interests. The neo-institutional analysis, on the contrary, takes into consideration the role of transaction costs, i.e. the costs of making the legal-economic system effective. Even here the firm is still seen as a *nexus of contracts*, somewhat framed by *contractual* institutions, notwithstanding the fact that the background analyzed is larger, and the impossibility of drafting a complete contract is taken into consideration.[15]

The firm cannot be a mere nexus of contracts, as these are inadequate instruments to describe a legal-economic reality so complex in itself, and dynamic by nature (Biondi, this volume). Contracts simply assess the instantaneous unanimity among decisional individual units. In addition, the peculiar activity of the firm – production – is performed only partially through contracts.[16] Even though some recent theoretical developments focus on the "contract incompleteness," they still believe these are contracts in nature, while Biondi and Moore and Rebérioux claim that "off contract" matters are constitutive of every social relation or body, such as firms. The scope of contractual analysis therefore tends, by definition, to shrink.

This volume has introduced and defined the basic concept of the firm-entity as a whole system where all its distinct elements are coordinated and based on fiduciary responsibilities and bound rationality. It follows that the exclusively formal view that defines an organization as a nexus of contracts appears to be restrictive, and does not take into consideration the features of "substance" which characterize every economic organization. The contractual approach acts, more or less implicitly, on the belief that contractual agreements can, in all cases, provide a "complete" regulation of every specific position of interests. In addition, it works from the standpoint that, once the contract is signed, there is no further possibility for related parties but to specify a set of dispositions within the given boundaries (i.e. its clauses). From this contractual viewpoint, both incomplete contracts and completing institutions (such as the board of directors or residual controlling rights) are contractual in nature. This theoretical and heuristic construction may represent the world of exchange contracts, but does not appear to represent the reality of the firm's legal-economic system with all the multilateral agreements *with a common scope*. These institutional agreements are drafted in order to regulate the common activity to be performed, and have

the special feature of requiring, for their performance, an indefinite series of new juridical bilateral or multilateral acts. Therefore, on the one hand, they cannot and will not regulate the shareholders' interests in a comprehensive way; on the other hand, they constitute a complex and dynamic institution whose goal is the regulation of the performance of the whole activity.

Finally, when the economic substance of the firm is seen simply as a nexus of exchanges where the sole legal forms considered are contracts, all the implications related to "economic development" are lost (point (iii)). Contracts are agreed at certain times under certain conditions, and are not able to take into account the complexity and dynamics of production, incomes and results. The agreement between the parties of a contract appears simply as an internal and particular moment within the whole process of development of the firm-entity.[17] Under the economic nature of the firm as an entity, contracts agreed with and within the firm-institution are only elements of a wider and more complex system of coordination in which they find their institutional framework. Contracts are (and deserve the right to be) a part of the firm, but they are far from constituting the whole firm or its whole legal-institutional structure. For such reasons, we agree that a contractual analysis has its own role, but we claim that a more comprehensive, synthetic and fruitful approach is required to understand the firm as a real entity.

## *The firm as the object of property rights*

The theory of property rights (Hart 1995; Hart and Moore 1990) states that the firm conceived as a company is just a portion of the wealth of its legal owners – the shareholders – whose role is that of owners of the assets as well as of the firm. They are entitled to make all the decisions about the use of these assets except those which, according to the law or the contract, are reserved for third parties: in other words owners are granted the "residual rights of control." In this way "property rights" can be interpreted as a mechanism to fill the gaps of incomplete contracts.[18]

According to this theory, property rights are justified since shareholders' investments are completely specific to the enterprise risk, while other stakeholders are granted exit options: employees, for example, besides the protection established by contractual clauses, may find a new job thanks to their cumulated competences and capabilities.

In fact, the limited liability company – one of the legal forms of the entity-firm – was set up in the first half of the sixteenth century, in order to limit the huge risks involved in oceanic trade, and developed when industrialization required capital concentration and risk sharing. In particular, it gained ground with the Napoleonic Code (at the beginning of the nineteenth century) and – from the financial point of view – mainly with the industrial groups of the second industrial revolution at the end of the nineteenth century. All kinds of "company" find their modern origin in the liberal ideology of protection and fiduciary conferring of duties, which, by nature and historical evolution, lies at the meeting-point of

public regulation and private action that needs to be both regulated and encouraged.

The characteristics of the limited liability company are well known: (i) the company is liable only within the amount of its equity; equity apart, shareholders are not liable for the corporate obligations, nor are directors; (ii) shareholders' identity is irrelevant, both as regards reciprocal relations and toward the entity itself; (iii) contributions to corporate equity are divided in fungible shares representing freely negotiable credit instruments; (iv) shares, historically disposable with some difficulty, have now become liquid on securities markets and stock exchanges. Berle (1965, reprinted in this volume) already recognized that *exit* is the special feature of the modern shareholding, as shares can be sold on the stock market. Nevertheless, shareholders can ask for special contractual protection, just like the other stakeholders, in order to secure their investments. In any case, shareholders are not the only constituency to make specific investments when they enter into relation with the company: for example, the investment in human and social capital made by employees of a company is as specific as that made by shareholders. As Mattei (1997: 38–39) has pointed out, to recognize the company's "veil" between shareholders and the outside world introduces into economic analysis a "real paradigm shift from the traditional natural law model assumed by economics, because it incorporates liability, an idea traditionally considered antithetical to that of property."

The featuring point of the property rights view is therefore the special and unbreakable link between the firm, conceived as a legal form, i.e. the company, and its shareholders in their capacity of owners of that firm. As Gynther (1967: 27) stated:

> Those who hold the proprietary concept perceive the firm as being owned by a sole proprietor, a set of partners, or a number of shareholders. The firm's assets are looked upon as being the property of these people and the liabilities of the firm are their liabilities.[19] [ . . . ] The proprietors are the center of interest at all times [ . . . ]. Profits are perceived to be the property of the proprietors (and not the firm) at the time they are earned, whether they are distributed or not.

According to this view, the legal entity of the company is merely a device of a representative nature by which the business in association may be conveniently administered with well defined legal privileges framed by legal dispositions.[20]

As Berle (reprinted in this volume) pointed out, legal evolution has made such a concept of ownership obsolete, especially in the way assumed by the theories of property rights: assets are now the property of the firm, ownership consists (i) of the right to dividends (under the discretion of the management as a matter of fact), and (ii) of the right to residual assets after liquidation.

From the legal point of view, once the shareholders have granted the cash-capital and/or the assets to the company, that capital or assets becomes the property of the company. As a conclusion, shareholders are simply providers of

financial resources different from others since they are not compensated through interests, but through dividends.[21] The company, therefore, responds to investors, not to owners.[22] Assets belong to the company just as inventions and innovations realized by employees – for instance patents – become the property of the company.

Furthermore, the firm as an entity gives life to relations between constituencies, including shareholders, different and broader than a set of contracts, complete or not. Under the legal-economic system of the firm, companies are special institutional devices that may enter into transactions independently of the identity of individual shareholders as well as of creditors and other lenders, since none of these subjects is entitled to "own" the firm. Even Blair (2001: 3) stated that:

> shareholders do not have the set of rights and responsibilities that we associate with "ownership" in other contexts. Shareholders do not directly own any of the assets used in production, nor do they own the output of the firm. They do not directly make any of the decisions about the use of the assets. They do not even bear full liability for the activities of the firm.

Kirat (1993) argued against the enterprise theory based on "ownership," and introduced two types of property rights to challenge such an old-fashioned framework. The first provides the owner with an absolute right to exploit his own assets against any other party, the second one represents only a relative right derived from the responsibility involved by the utilization of limited resources. The firm-entity approach obviously recognizes only the relative right, since the absolute right is inconsistent with the positive law that regulates the *libertas sub lege* (freedom under the law) granted to economic agents. As Mattei (1997: 38) further pointed out:

> According to Hegel, collective interest is something ontologically different from the sum of the individual interests: private property, therefore, is not an institution recognized as promoting social progress but merely a means of guaranteeing expression of individual free will. In this conception originated both the idea of inherent selfishness in the exercise of property rights and the need for public law regulation to limit such selfish exercise.

The legal and logical distinction between the firm, the company and the shareholders can in no way be considered a starting point, but inversely as the main conclusion explaining and justifying the well known separation between property and control, shareholders and managers. On this topic, some seventy-five years ago, Demaria (1930: 255) pointed out that:

> the fundamentals of classicism and marginalism lie in the assumption of a fixity of social institutions, but the most important of these ones, ownership, is not immobile in its social function, notwithstanding (and maybe

due to) the special technique of large companies which allow their managers to pursue particularistic interests against those of the administered subjects.

In conclusion, from the legal-economic point of view, shareholders own merely their own shares, and cannot be defined as "owners" of the firm, nor can managers. Managers have fiduciary duties toward shareholders and stakeholders in general,[23] and they are not allowed to freely and discretionally dispose of all the assets of the company. The conclusion is that the firm as an entity has no "owners," because no constituency has the totality of rights and responsibilities which pertain to the institute of "ownership."

### *The firm as a going economic coordination*

As Commons (1924: 172) claimed, "if such a thing as a going concern actually exists, distinguishable from physical things, then failure to recognize it perpetrates injustice."

According to the comprehensive approach developed by this volume, the entity-firm is interpreted as a dynamic and holistic system performing an economic activity. And this is true apart from the legal forms it adopts, since "it is the combination of the parts that grants them purpose and meaning, thereby transforming and making them components of a whole."[24] Drawing upon Commons and Zappa, the firm is thus understood as a system of coordinations framing and enhancing human activity. It becomes an "institution," i.e. a "collective action in control, liberation and expansion of individual action"[25] able to prompt, frame, and enhance individual economic capacities.[26]

This legal-economic system expressed by the firm-institution is based on the encounter between management discretion and legal guidance. This system establishes a relation between the economic substance of the firm and the state legal system, and focuses on the overwhelming influence that legislative decisions exert upon the production and allocation of wealth.

A firm's coordination is established either by the management or is semi-spontaneous, and in any case suited to dispositions of the top-down type (the law) or the bottom-up type (self-regulation). In either case, within the legal-economic system of the firm, specific internal norms complement the law: this concerns the (often informal) regulating codes that factually govern the day-by-day functioning of the firm's coordinations. These norms tend to reinforce the governing structure, in some cases complementing, in some cases overriding the written dispositions. We are clearly speaking of extra-contractual norms, which play such an important part in the life of stable and lasting organizations: they go on working independently of the change of their members.

We must therefore recognize that, in such a way, the firm expresses its own institutional role of law-co-maker by coupling its own self-regulation (mainly through by-laws) with the State legal system. This role has a historical imprint, since – under both common and civil law – the State can not fully regulate the

enforcement of its laws once and for all: it must make space for many other institutions too.

The firm being a dynamic institution, never turns into a single or single-type governance structure. Neither written contracts nor other law-based institutional devices are self-standing: they are shaped by the entity-firm as a law-co-maker, as a social whole with its own living reality. To treat the firm as an institution means a comprehensive unity, whilst contracts express a multiplicity of parties. The mere existence of subjects connected to each other by transient interactions is not sufficient: the firm requires the stable and coherent connections proper to a social structure, in which formal organizational structures and functions identify the roles of subjects and also qualify them. In any case, the juridical forms we are speaking of only partially express the economic substance of the enterprises. The legal form represents mutual obligations, while the economic substance, on the contrary, evolves with the coordinations carried out to fulfill the common interests of their constituencies.

Such a dynamic system of the entity-firm can be in fact represented by a pair of joint concepts, as it realizes at the same time (i) unitary multiplicity, and (ii) stable mutability.[27]

Unitary multiplicity refers to the enterprise as a system of capitals and incomes, resources, prices, investments and financings, decisions, risks, persons, structures and values. This system is an economic complex extended in time and space and its elements are linked by dynamic relations of complementarity, interdependency and interaction. Its stable mutability refers to the ongoing existence of the relations which, with the passing of time, hold together all the coordinations of the system notwithstanding its changing elements.

## Some consequences of the firm as a legal-economic system

### *The industrial group as a unitary enterprise*

The industrial group is a coordination of productive units, juridically autonomous, connected through a relatively stable and coherent organization in order to best pursue the goals of such a complex system. The unitary economic coordination and the formal autonomy of the companies participating in the group constitute qualifying features of this institute. By means of its unitary coordination, the group achieves the scope to realize unified vision and actions for all the controlled companies and, more generally, for all the activities, whatever legal form they involve (see also Gindis, this volume). Through formal autonomy of components, the group succeeds in giving structural flexibility and delimitations of risks and obligations to the economic body, unitarily considered.

Historically the legal systems adopted the atomistic approach of corporate functioning and of the autonomy of the companies as separate decision centers.

On the contrary, economic practice generated many instruments to coordinate subsidiaries. The group can be created by means of (i) the incorporation of subsidiaries; (ii) the stipulation of special agreements to give life to contractual

groups; (iii) the stipulation of other agreements to bind (or to reinforce the binding of) companies (franchising, supplying of technology, trade mark licensing, assignment of know-how and financing agreements, among others); and (iv) connections between the persons in charge of the management of the different companies.

The very economic identity of industrial groups is an articulate issue taking into consideration the individualization of both the parent company and a cluster of limited liability companies; the individuation of its operative characteristics; the analysis of the degree of autonomy of the single subsidiary (given its juridical personality and formal independence).

Industrial groups display, at the same time, both unity and plurality: more companies find their own purpose in a unitary firm, which is run in a mediate way by the holding and in a direct way by the controlled corporations.[28] The plurality of juridically autonomous companies – subsidiaries – leads to economic unity by the strategic management of the holding.

From the point of view of the enterprise, each subsidiary has a partial business purpose, i.e. either a specific sector of the whole activity performed by the group, or a territorial competence, or a particular phase of the productive process, in any case giving rise to the division of the two core activities of the firm: in short, "management" and "production."

This group of juridically differentiated companies represents a unitary firm for a multiplicity of components and combinations: a unitary entity finances the single companies; a unitary entity is remunerated by the activity of the single companies, because the revenues realized by them flow, directly or thorough sub-holdings, to the parent company; a unitary decisional center, the same, manages the whole group, having directly or indirectly the control of all companies (sometimes through the totality or the majority of the votes in any of them).[29] To summarize, with a group we do not have a firm for each company, but an unitary firm: the holding as a whole operates through its subsidiaries as units. The dynamic accounting principles – discussed by Biondi in this volume – recognize this factual reality both by consolidating the firm as an entity according to the limits of risk and control, as well as by somewhat disregarding mere formal frontiers established by legal arrangements.

Among jurists, on the contrary, it is current wisdom to understand the group, as taught by the law, as the mere common plurality of companies which compose it. The unity of the group as a whole, being merely economical, is a logically superior controlling reality, absolutely irrelevant to the law (except for the special legal dispositions that, in specific cases, give unitary relevance to the group).

Notwithstanding the formal atomistic approach generally adopted by the law, in the last few years jurisprudence has begun to consider the industrial group in its real economic nature of unitary firm,[30] developing (i) the theory of "unitary business" for taxation – in order to take into account the activities carried out worldwide by the group through all its subsidiaries – and (ii) the theory of "integrated enterprise" for labor law – with the aim of identifying within the

group the unitary employer independently from the single company that signed the labor contract with the employees.

Jurisprudence has then identified some criteria for the identification of a unitary firm within a group of formally individual companies, such as (i) the same organizational and productive structure; (ii) integration between the activities performed by the various companies of the group to pursue a common interest; and (iii) the technical and administrative-financial coordination giving the industrial group a unitary scope and intents that override the companies.

The group *factually* represents the typical institutional structure of the medium- and large-sized firm of the twentieth century, created with the goal of limiting the risk of different coordinations, but in some cases established with the aim of bypassing relevant or special regulations of company law (so-called "Chinese boxes").

In any case, if the firm is a unitary system, and therefore – as discussed above – a unitary juridical system, it becomes necessary to recognize for all the components a unitary liability borne by the holding.

As far as liability is concerned, the fact that a subsidiary has a separate legal personality is not sufficient to exclude the possibility of imputing its conduct to the parent company. This case occurs in particular when the subsidiary, although having a separate legal personality, does not decide independently upon its own conduct on the market, but carries out, in all material respects, the instructions given by the parent company.

Whenever the subsidiary does not enjoy any real autonomy in determining its course of action, the legal dispositions may be considered to be inapplicable in the relationship between it and the parent company with which it forms one economic entity.

So, in view of the unity of the whole group, the actions of the subsidiaries may in certain circumstances be attributed to the parent company. In particular, the holding which, exercising a dominant influence in managing and coordinating activities, acts in its (or others') entrepreneurial interest against the principle of "fair" management of the companies which are part of the group, should be directly liable at least: (1) to their shareholders for the damage caused to profitability or to the value of the corporate participation; (2) to the corporate creditors for the damage caused to the integrity of the corporate patrimony; and (3) finally to the employees for the social investments made by them in the firm itself.[31]

## ***The common interest of the constituencies in the firm***

The expression "corporate interest" means the interest which can be considered intrinsic to the company as an entity. Corporate interest (a) constitutes the parameter for the behavior of directors who have to make their decisions according to the pursuit of this interest, and (b) limits the power of the majority, whose decisions, taken for its own interest and against the corporate interest, may be invalidated. In other words the corporate interest is the goal which leads

the corporate organization to act in order to make resources economically productive. As stated by Berle and Means (1932: 293):

> The shifting relationship of property and enterprise . . . raises in sharp relief certain legal, economic, and social questions which must now be squarely faced. Of these the greatest is the question in whose interests should the great quasi-public corporations . . . be operated.

Nevertheless, our view of the business entity as the legal-economic system over comes the corporation as the preeminent unit of analysis. The *form-based* notion of corporate interest has to be rediscovered and understood in the *substance-based* notion of common interest related to the whole firm as a collective, evolving undertaking.

With the word "interest" we mean the relation between a subject that has a need and the activity capable of satisfying such a need: a relation which is at least partly recognized by the general and abstract provision of a norm (Jaeger 1964). In particular, the firm has a joint and collective nature, and it is a system formed by a complex net of coordinated and several interests.

The expression "conflict of interests" is a well known one, and it depicts the fact that different subjects have an identical need which is satisfied by the same good when this good can not totally or partially satisfy the needs of all these subjects. A collective interest can be described, on the contrary, as the interest of individual subjects in the light of the interest of the other subjects and of the common purposes they want to achieve and in which they are taking part, as well as in the light of the interest of the society and of the human community to which both the subjects and their economic activity belong.[32]

In the juridical doctrine the issue of the corporate interest has been profoundly discussed by both contractual[33] and institutional approaches, or, with a more modern vocabulary, from the viewpoints of both "shareholder value" and "stakeholder value."

In international literature the debate is historically related by two papers published in the *Harvard Economic Review* in the early 1930s by Berle and Dodd.

Berle (1931: 1049) affirmed that: "all powers granted to a corporation or the management of a corporation . . . are necessarily and at all times exercisable only for the rateable benefit of all the shareholders as their interests appear." This is a shareholder-centered model of fiduciary duties.

One year later Dodd replied to this assumption, declaring his opposition and underlining the limits of management which focuses exclusively on the pursuit of maximum shareholders' profit. On the contrary, managers should be "guardians of all the interests which the corporation affects and not merely servants of its absentee owners" (Dodd, 1932: 1157), due to the fact "public opinion, which ultimately makes law, has made and is today making substantial strides in the direction of a view of the business corporation as an economic institution which has a social service as well as a profit-making function" (ibid.: 1148). Recalling this debate, it has to be underlined that, more than twenty years later, Berle (1954: 169) accepted the point, finally conceding to Dodd.

The contractual theory conceives the corporate interest as coincident with the exclusive interest of shareholders, no other constituency being considered. Nevertheless, once the firm is interpreted as ongoing economic coordination, we must take into consideration not an egoistic interest of the firm itself, but a collective one, or even better, more joint and collective interests, which in turn create a binding solidarity among the subjects with such interests. This occurs when the coordination among the interested constituencies allows the pursuing of the business purpose and, at the same time, the fulfillment of their needs: a solitary constituency could not achieve the same outcome on its own. In other words, solidarity among different interests consists of a relation of complementarity and mutual instrumentality of different subjects, considering that the interest of just one constituency cannot be satisfied if the interests pertaining to the others have not been jointly accomplished. Therefore, the holders of the corporate interest cannot be the shareholders alone.

So, a revision of the contractual approach seems to be called for: if the firm is a system, the corporate interest should pertain to the system as such. We are not dealing with the mere aggregation of interests, but – the firm being a going economic coordination – with a coordination of interests called "common interest."

As dominant groups of shareholders can betray such common interest,[34] managers should adopt their resolutions, taking into account not only the "egoistic" interests connected with profit maximization (shareholder value), but also the interests of other corporate groups (employees, creditors, others), increasingly widespread interests (pollution, energy-saving, and so forth) and the general context (respect of human rights, protection of infants and minorities), as already discussed by Avi-Yonah and Sivan in this volume.[35]

In order to pursue such institutional goals, at least two routes are viable: (1) the dominant constituencies of the firm may include social goals within the interest pursued by the firm, for example by the appointment of representatives of different stakeholders in the corporate bodies, and (2) the legal system may impose on the firms themselves to maintain socially oriented behavior by means of formal legal provisions.[36]

Therefore, the common interest pursued by firm-entities, recognized and enforced by the state legal system, must be led by the integration of the social function with private ownership (Zappa 1956/57: 29) and not by the impracticable "shareholder value maximization" championed by the defenders of the equilibrium-framed theories.

The definition of common interest and its contents are not the issue of a "market machinery," but on the contrary the result of the intervention of subjects and bodies belonging to the power structure: in principle this means management with its governing processes but – in the case of arbitration – also judges applying their own discretionary power.[37] Moreover, some interests of individuals and groups may be selected by the state legal system (*policy of law*) and defended insofar as they are compatible with the social dynamics.[38]

The firm – as Simon stated – has neither the information nor the analytical capabilities to pursue shareholder value maximization (see also Anthony 1960,

reprinted in this volume). On the contrary, the common interest can be only pursued by means of day-to-day management and, if necessary, arbitration. Such dynamic management is far from any instantaneous equilibrium. On the contrary it conforms with time to every change in the legal system (new laws), in the same way as it adapts to changes in prices, products, and technologies: the only unchangeable law is the law of never-ending coping.[39]

## Conclusion: the legal-economic system of the firm between external regulation and internal governance

The relevant institutional role of the existing legal system may be interpreted as a negation of the specific concepts of the law and economics school, according to which firms should act freely in a system of perfect competition (free from bounds set up by juridical and social institutes), leaving to the market alone the task of selecting the "most efficient" firms. According to such assumptions, judge Posner supported the idea that corporate bankruptcies demonstrate the existence of an efficient market able to expel inefficient companies.[40] But what about the costs for shareholders, for constituencies in general, for the community as a whole?

In addition, the presence of corporate groups can be observed, whose control is not freely in the hands of the market as they are in fact dominated by the aforesaid alliance of management and dominant shareholders who have a vested interest conflicting with minority shareholders and other stakeholders. This fact gives prominence to issues which are difficult to be solved effectively by the market in general and by the "market for control" (i.e. takeovers) in particular.

We must recognize that self-regulation and the market are not alone: not only does the state with the government of corporate laws integrate the two alternative poles of the market and private ordering, traditionally considered as alternatives, but its regulatory intervention also interacts with the autonomous regulation provided by the market.

Once the market cannot work properly, or becomes inadequate for producing effective results, state intervention occurs, by means of mechanisms that are different in nature if compared to the market itself. The allegation that such state intervention (laws or dispositions enacted by public bodies) is never intrusive and does not distort the operation of the market is contrasted here by the recognition that (i) the market itself is not an autonomous reality, and (ii) juridical arrangements appear to play a fundamental role in its existence.

The legal-economic system of the firm is not set, nor is it a fixed structure, but an evolving reality, constantly under the interplay of management, courts, and government, which may change and materially modify the economic framework, so changing the meaning of contracts and business activities.

In this perspective, the market is not the sole and general regulatory system of economic production and distribution: the management as such, and sometimes the courts (and the government) acquire "authoritative faculty of political economy" (in Commons' words). It follows that it is not the market alone, but with

management, government and the judges, which establishes complex and dynamic economic processes of production and allocation, and is utterly different from the Walrasian "neutral auctioneer." Such processes are clearly not mathematical, but economic and social in nature, and thus they are better understood through an economic-institutional understanding, than by mimicking physical models.

The actors we have mentioned shape the economic and monetary processes of fulfilling human needs. For instance, as Commons early recognized, judges – by means of doctrinal interpretations of legal dispositions (civil law) or judicial precedents (common law) – can influence the allocation of economic resources and finalize the legal-economic system of the firm towards economic policy goals. In parallel, they can act in a different way from the logic of the market-price system, also making the common interest of the firm effective.

Continuing with Commons, judges' decisions may evaluate the actions of managers as fiduciaries, interpreting the "common purpose" involved in the separation between ownership and control. In this perspective, directors are entitled to manage the conflicts they play a relevant part in (for instance, the payment of the dividend is a cost for the company while it is a profit for the shareholder). Thus, the leading idea of common interest becomes fundamental, as Zappa (1956/57: 60) clearly stressed:

> In firms the necessity is present of a vivid solidarity between all the factors operating for the production and everybody sees that, for a bigger production, for a more diffuse consumption and for a congruous distribution of the incomes between all the productive factors, it would be highly opportune to have a new legal system which could effectively embank the excessive desire for profitability which moves the subjects who pro tempore exercise the *highest* control over the production of the firm.

We are speaking of such problems as the common interest, and its effective synthesis and application. As far as this problem is concerned, more than a century ago Pescatore (1879: 24) offered us the choice between two principles which still serve as a reference:

> First principle: limited liability companies, as all the other categories of companies, civil or commercial, are a contract of private law: the freedom of the agreements, and the free exercise of any property right legally acquired provide the rules, the first for the constitution, the second for the by-laws in the administration of the limited liability company. Second principle (opposed to the first one): limited liability companies are institutes of public law; their promoters perform a real social function, subject to the bonds of the legal dispositions regulating the social functions, and when they are incorporated, their administration is a government of common interest, to be regulated by the guarantees, which are proper to the government of common interests.

We assume that the firm is of a private nature, since it arises from individual economic activity. Nevertheless, as it is at the same time an institution, its legal-economic system deals with interests that are public in nature.

Once the overriding reality of the legal-economic system of the firm as an entity is considered, distinguishing between inner interests which are either (i) more easily defensible by external ties, or (ii) whose protection can be improved through the creation of specific internal working norms, concerns the governance of the firm – both self-regulated or legally enforced – treated by Moore and Rebérioux in this volume. The same goes for the establishment of the often subtle boundary between external interests (environment, protection of human rights) and those able to find representation by and within the corporate bodies, as the German model of governance (*Betriebsverfassungsgesetz*) has already suggested.

## Acknowledgment

The author thanks A. Canziani and Y. Biondi for their helpful comments.

## Notes

1 The discordance between economic substance and legal form of an operation often derives from the different speed of the economy in comparison to the law that regulates it. In particular the rapidity of development of financial markets has forced the operators to adapt the existing juridical instruments to new financial instruments, so creating a whole disarray between pursued economic goals and utilized contracts. For such reason accounting has always considered the economic substance of the operation as the founding element for its recognition and its report in financial statements, in order to assure the well known "true and fair view."

2 See Zappa (1956/57). See also Kirat and Gindis, this volume.

3 Fama and Jensen (1983); Easterbrook and Fischel (1991).

4 According to more relevant doctrine, the deed of incorporation can be regarded either as a recognition (real entity view), or as a right (aggregate entity view), or as a privilege (artificial entity view).

5 See Machlup (1967: 9):

> To confuse the firm as a theoretical construct with the firm as an empirical concept, that is, to confuse a heuristic fiction with a real organisation like General motor or Atlantic & Pacific, is to commit the "fallacy of misplaced concreteness." This fallacy consists in using theoretic symbols as though they have a direct, observable, concrete meaning.

In fact, Machlup was advocating the legal fiction view *vis-à-vis* the legal person view, two perspectives we are considering here as compatible and complementary in the artificial person view.

6 In the medium-large companies the dimension of the solutions and organizational structures is relevant with reference to the whole cluster of the controlled companies and the relevant strategic decisions. These decisions are an expression of the real power of influence and control, and are taken by the top management of the parent company. See Berle (1947: 345):

> the courts have either (a) erected corporate personality which the State had not granted; or (b) disregarded corporate personality where the State had granted it,

> for the real purpose (in both categories) of giving legal effect to factual relationships set up between an economic entity and an outsider.

7 According to the special characters Zappa (1956/57) attributes to a social institution, see Biondi (2005: 114).
8 See Canziani (2004).
9 It is worth recalling here the example of the accounting system. It finds its *raison d'être* inside the firm (structure of the accounts, accounting for management), but at the same time outside the firm (accounting rules issued either by instituted standards-setters such as IASB, or self-regulated by the profession itself). The two reasons interact in creating an accounting "legal-economic system" – i.e. the accounting system becomes a constitutive part of the legal-economic system that is the entity-firm – and in defining its role in the activity of the firm. The accounting system so created by the firm acquires a legal meaning and, by the disclosed financial statements, a legal form as well.
10 See Trustees of Dartmouth College vs. Woodward, 17 U.S. (4 Wheat.) 518, 636 (1819). As a contrary example we have taken into consideration a liquidated company which, even after the winding up of its assets, continues the previous activity of production.
11 See Berle (1947, this volume). For example in some state legal systems, trade unions lack the juridical personality but they are in any case fully allowed to enter into collective contracts granted with general effectiveness.
12 From the legalistic point of view of the "company," instead, the firm's legal-economic coordination comes to something considerably different: by one side, the coordination will be fragmented and reduced to many heterogeneous relations which take place between the persons who interact; by another side, as legal unit, the company will be considered as an *universitas rerum*, i.e. as a mere aggregation of assets, obligations and claims pertaining to the legal person that the company is.
13 See Troisi (2004).
14 For the role of directors as agents or trustees, see Blair and Stout (1999).
15 See Di Laurea (2001).
16 Already in 1908, Mitchell showed that the employee's wage did not result by a contract but from a complex procedure involving institutional matters. A bilateral bargaining, thus, is never enough to determine the wage payed.
17 See Leo (1994: 11ff).
18 Contracts are incomplete due to the limited rationality of the actors and for the incompleteness of the information provided to them.
19 See Sprouse (1957: 370b): "shareholders … are the owners of the corporate assets and obligors of the corporate debts."
20 See also Sprouse (1957: 370b) on this contractual mixture between the aggregate and artificial person view of the company.
21 See Littleton 1953; Raby 1959.
22 It should be remembered that in the event of winding up, shareholders will be the last ones to be satisfied.
23 See Donaldson and Preston (1995).
24 Raby (1959: 452b). Further developments in Biondi, this volume.
25 See Commons (1931: 651).
26 See note 7 and accompanying text.
27 According to the Italian tradition presented by Canziani (this volume).
28 See Galgano (2005).
29 See Galgano (2002).
30 According to Berle (1947): reprinted in this volume:

> in an apparently growing number of situations the courts, in effect, mold the corporate situation to the economic fact; that the economic fact is the actual

> business enterprise as carried on by the component individuals active in it; and that the entity, with its attendant consequences of a particular body of assets and operations, is then given legal attributes which would have been given to it had it been a body corporate duly and properly enfranchised by the state.

31 See Blumberg (1990: 69):

> It is time to consider whether the parent and affiliated companies of the group should also be liable in the particular case for the duties and obligations of the relevant subsidiary in order either to protect persons dealing with companies of the group in case arising at common law or to implement governmental controls and prevent their frustration and evasion more effectively in case involving statutory law.

32 A special kind of collective interest is represented by the interest of each constituency. It is then necessary to analyze if such constituency's interests are "of group" (for example the stable shareholders of a small limited liability company whose shares are not listed), also from a timing point of view, or "of series" (for example all the potential shareholders of a listed limited liability company).

33 See Jensen (2001).

34 Moore and Rebérioux (this volume) argue that the controlling coalition between managers and dominant shareholders, rather than gatekeeper failure, is directly responsible for the failures and shortcomings of the Enron era.

35 Already François Perroux and the German school of the "Unternehmen an sich" [whole enterprise as such], fostered by Walther Rathenau, warned us that medium-sized and large firms should depart from a strict shareholders-owners view (with their private interests), in order to assume their own proper relevance as active parties of the economic system. The firm conceived as a public-oriented institution should provide incomes and products for the community, create jobs, improve technologies and facilitate scientific progress.

36 See Denozza (2005).

37 In this context, it is useful to analyze the goals of accounting from the proprietary (shareholder) or the entity (stakeholder) perspective. For the entity theory, Littleton (1953: 34) underscores that:

> The public accountant has an important moderator-function to perform . . . : he has the complex task of fashioning complex materials into a form that is useful to a mixed community of interests; he has the important social task of holding his independence in such high esteem that his report can be accepted by all parties.

The opposite proprietary approach was taken by the "proprietorship theory" in the United States at the beginning of the century. For example, Sprague (1907: 59) claimed that: "the whole purpose of the business struggle is the increase of wealth, that is increase of the proprietorship . . . the all important purpose of the proprietary accounts is to measure the success or failure in increasing wealth."

38 See Leader (1995: 90):

> the independence of the company's interest does not consist in its being unattached to the interests of anyone, but rather in its being attached to some of the interests of all of those affected by it; that is, those interests which, when you satisfy them, help you to further the purpose for which the company exists.

39 Drawing upon the cognitive and epistemic structure of production, of which Machlup (1946) was a precursor and Simon (1960) the forerunner, later formalized by Cyert and March (1963), a further attempt to get the economic theory close to the social

reality is pursued, abandoning the hypothesis of full rationality of the economic operators and of perfect information.

40 The perfect legal system can be seen in harmony with the "pure theory of the law" by Kelsen: existence of complete and well defined rights with respect to any possible transaction for which an effective system of enforcement is available at zero cost.

## Bibliography


Alchian, A. A. and Demsetz, H. (1972) "Production, Information Costs, and Economic Organization," *American Economic Review*, 62: 777–95.

Anthony, N. R. (1960) "The Trouble With Profit Maximization," *Harvard Business Review*, 38 (6): 126–34, reprinted in this volume.

Beigel, R. (1900), *Das allgemeine deutsche Buchführungsrecht. Die kaufmännische Buchführung nach dem neuen Handelsgesetzbuch und der sonstigen neuen kaufmännischen Gesetzgebung; mit praktischen Beispielen, Tabellen und Formularien für den Handegebrauch bearbeitet* (The German Business Accounting Law), Leipzig: Verl. d. Handels-Akademie.

Berle, A. A. (1931) "Corporate Powers as Powers in Trust," *Harvard Law Review*, 44 (7): 1049–74.

——(1947) "The Theory of Enterprise Entity," *Columbia Law Review*, 47 (3): 343–58, reprinted in this volume.

——(1954) *The TwentiethCentury Capitalist Revolution*, New York: Harcourt Brace.

Berle, A. A. and Means, C. G. (1932) *The Modern Corporation and Private Property*, New York: Harcourt, Brace & World.

Biondi, Y. (2005) "G. Zappa, T. Veblen, J. R. Commons, e l'impresa come istituzione economica" [Zappa, Veblen, Commons and the firm as an economic istitution], *Storia del Pensiero Economico*, 1: 93–120.

——(2006) "Accounting and theEconomic Analysis of the Firm as an Entity," in M. Dietrich (ed.) *The Economics of the Firm: Analysis, Evolution and History*, New York and London: Routledge.

Blair, M. M. (2001) "Team Production Theory and Corporate Law," *Working Paper Series in Business, Economics, and Regulatory Law and Public Law and Legal Theory*, paper no. 281818, Georgetown University Law Center. Available at http://papers.ssrn.com/abstract=281818

Blumberg, P. I. (1990) "The Corporate Personality in American Law: a Summary Review," *The American Journal of Comparative Law*, 38 Supplement., *U.S. Law in an Era of Democratization*: 49–69.

Canziani, A. (2001) "I processi competitivi fra economia e diritto" [Competitive processes between economics and law], University of Brescia, Department of Business Administration, Working Paper no. 15, September.

——(2004) "La natura economica dell'impresa" [The economic nature of the firm], University of Brescia, Department of Business Administration, Working Paper no. 32, June.

Coase, H. R. (1937) "The Nature of the Firm," *Economica*, n.s., no. 4, November: 386–405.

Commons, J. R. (1924) *Legal Foundations of Capitalism*, New York: Macmillan; reprinted New York: Transaction Publishers, 1996.

——(1931) "Institutional Economics," *American Economics Review*, XXI: 648–57.

Cyert, R. M. and March, J. G. (1963) *A Behavioral Theory of the Firm*, Englewood Cliffs, N.J.: Prentice-Hall.

Demaria, G. (1930) "Saggio sugli studi di dinamica economica" [Essay on studies of economic dynamics], *Rivista internazionale di scienze sociali e discipline ausiliarie*, XXXVIII, 3: 222–57.

Denozza, F. (2005) "Interesse sociale e responsabilità sociale dell'impresa" [Social interest and responsibility of the firm], in L. Sacconi (ed.) *Guida critica alla responsabilità sociale e al governo d'impresa*, Rome: Bancaria Editrice.

Di Laurea, D. (2001) "Complessità istituzionale, capacità ed eco-diversità del sistema produttivo" [Institutional complexity, capability and eco-diversity of the productive system], *University of Siena, Department of Political Economics*, Working Paper no. 324, July.

Dodd, E. (1932) "For Whom Are Corporate Managers Trustees?," *Harvard Law Review*, 45: 1145–63.

Donaldson, T. and Preston, L. (1995) "The Stakeholder Theory of the Corporation: Concepts, Evidence, and Implications," *Academy of Management Review*, 20 (1): 85–91.

Easterbrook, F. H. and Fischel, D. R. (1991) *The Economic Structure of Corporate Law*, Cambridge, Mass.: Harvard University Press.

Fama, E. F. and Jensen, M. C. (1983) "Separation of Ownership and Control," *Journal of Law and Economics*, 26 June: 301–25.

Galgano, F. (2002) "I Gruppi di Società" [Corporate groups], in P. Zanelli (ed.) *Società, Contratti, Metodo*, Milan: Giuffrè.

——(2005) *Direzione e coordinamento di società* [Corporate strategic management and coordination], Bologna: Zanichelli.

Gynther, S. R. (1967) "Accounting Concepts and Behavioral Hypotheses," *The Accounting Review*, XLII (2): 274–90.

Hart, O. D. (1995) *Firms, Contracts, and Financial Structure*, Oxford: Clarendon Press.

Hart, O. D. and Moore, J. (1990) "Property Rights and the Nature of the Firm," *Journal of Political Economy*, 98: 1119–58.

Jaeger, P. G. (1964) *L'interesse sociale* [The corporate interest], Milan: Giuffré.

Jensen, M. C. (2001) "Value Maximization, Stakeholder Theory, and the Corporate Objective Function," *Journal of Applied Corporate Finance*, 14 (3): 8–21.

Kirat, T. (1993) "De l'analyse économique du droit à de nouvelles alliances" [From the economic analysis of law to new alliances], *Revue Economique*, 1057–87.

Leader, S. (1995) "Private Property and Corporate Governance. Part I: Defining the Interests," in F. Pacfield (ed.) *Perspectives on Company Law*: Amsterdam: Kluwer Law International.

Leo, E. M. (1994) *Forma giuridica e materia economica* [Legal form and economic substance], Milan: Vita e Pensiero.

Littleton, A. C. (1953) *Structure of Accounting Theory*, Illinois: American Accounting Association.

Machlup, F. (1946) "Marginal Analysis and Empirical Research," *American Economic Review*, September: 519–54.

——(1967) "Theories of the Firm: Marginalist, Behavioral, Managerial," *American Economic Review*, 57 (March): 1–33.

Mattei, U. (1997) *Comparative Law and Economics*, Ann Arbor: University of Michigan Press.

Mitchell, W. C. (1908) *Gold Prices and Wages Under the Greenback Standard*, Berkeley: University of California Press.

Pescatore, M. (1879) *Filosofia e dottrine giuridiche* [Philosphy and legal doctrines], Turin: Fratelli Bocca.

Raby, L. Wn. (1959) "The Two Faces of Accounting," *The Accounting Review*, 34 (3): 452–61.

Romano, S. (1977) *L'ordinamento giuridico* [The legal order], Florence: G. C. Sansoni Editore Nuova.

Simon, H. A. (1960) *The New Science of Management Decision*, New York: Harper & Row.
Sprague, C. E. (1907) *The Philosophy of Accounts*, New York: Ronald Press.
Sprouse, T. R. (1957) "The Significance of the Concept of the Corporation in Accounting Analyses," *The Accounting Review*, XXXII (3): 369–78.
Troisi, R. (2004) "Teoria dell'impresa e responsabilità parapenale: le implicazioni organizzativo-gestionali," *Centro di Economia del Lavoro e di Politica Economica*, *University of Salerno*, Discussion Paper no. 89, December.
Williamson, O. E. (1975) *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: Free Press.
——(1985) *The Economic Institutions of Capitalism*, New York: Free Press.
——(1991) "Comparative Economic Organization: The Analysis of Discrete Structural Alternatives," *Administrative Science Quarterly*, 36 (June): 269–96.
Zappa, G. (1937) *Il reddito d'impresa* [The income to the firm], Milan: Giuffré.
——(1956/57) *Le Produzioni nell'Economia delle Imprese* [Production in the economy of firms], vol. 3, Milan: Giuffré.

# 17 Economics and finance of the firm as an entity

*Giuseppe Marzo*

## Introduction

Standard finance theory was born from the rib of neoclassical economics. Notwithstanding some assumptions have been modified or relaxed, this neoclassical finance theory (NFT) is still based upon: (1) methodological individualism; (2) the fully rational *homo economicus*; and (3) the capital markets' efficiency.

Despite its fast development as a comprehensive decision making theory, NFT is increasingly under attack because of the irrealism (or the anti-realism?) of those basic assumptions.

The main criticisms are usually addressed to both the fully rational agent and capital market efficiency, whilst methodological individualism is more tolerated. The rationale for this chapter is, on the contrary, that the evidence that the real world is a world of organizations and institutions, which cannot be reduced to the mere aggregation of single agents' behaviors. More dynamic and holistic features and implications have thus to enter the framework.

This chapter focuses on a finance theory built around the concept of a firm as an institution. Therefore it aims at shedding light on the effects that the introduction of the entity-firm concept has on finance theory.

In order to achieve its goal, the chapter is structured as follows. The second section synthesizes NFT from a methodological standpoint, also offering some insights on both its foundation and how it differs from so-called traditional finance. The third section outlines the problem of the missing entity-firm at the core of the analysis, and the fourth section presents the criticisms levelled against full-rationality and capital markets' efficiency assumptions. The fifth section offers an analysis of how the introduction of the entity-firm concept calls for renewing finance theory, and the sixth section addresses the topic of capital structure decisions under the light of the basic premises of the entity-firm-based finance theory. Finally, conclusions are outlined.

## The war of the worlds: traditional finance versus financial economics

Finance theory has been characterized by a long process of (r)evolution, with relevant shifts in methodology and in the analysis's focus. This section will

present this evolution from an historical and methodological perspective, and thus summarize the state of the art of "standard" neoclassical finance theory (NFT).

## *A brief history of finance theory from the methodological perspective*

Bearing in mind the articles currently appearing in the leading finance journals it could seem astonishing that before the 1950s finance theory was – when compared with and analyzed from today's finance theory standpoint – eminently descriptive, anecdotal and prescriptive. Generally speaking, it was also more focused on an institutional approach to financial matters and consisted in large part of ad hoc theories (Jensen and Smith 1984). At that time, major concerns were optimal investment, financing and dividend policies, working capital management and financial statements analysis. Researchers paid attention to the overall activities performed by the firm as well as to some specific financial topics. The methodology was eminently institutional and not based on complicated mathematical models build upon selected assumptions. As a matter of fact, very few analytic tools were used beyond accounting, arithmetic and algebra.

The first breakthrough was probably Miller and Modigliani's (1958) paper on the capital structure irrelevance,[1] even though the article by Markowitz (1952) on portfolio selection is often cited as the starting point of the new way to do finance.

Miller and Modigliani state that when markets are perfect and there is no taxation, capital structure (i.e. the ratio between debt and equity) is irrelevant, in the sense that it does not influence the firm's capital costs. As Stiglitz notes (1988: 121)

> it is ironic that a paper which purportedly established that one need not to pay any attention to financial structure – that financial structure was irrelevant – should have focused economists' attention on finance.

At a first glance, the importance of the article seems predominantly due to the fact that traditional approach to capital structure decisions was based upon the financial leverage mechanism under the assumption that – the debt being less costly than equity – a firm would be able to reduce the cost of capital by simply increasing its debt until an optimal level was reached. By questioning this current wisdom that an optimal capital structure always exists, Miller and Modigliani's paper stimulated an intense debate on the topic. Many interesting advances on capital structure derived from the attempts to react to and expand upon the Miller and Modigliani theorem.

Actually, the importance of the paper relies more fundamentally on the methodological approach it follows, which it is the one currently employed by finance theory.

First of all, the paper does not aim to propose a normative theory about capital structure, but a positive one. The paper shares, indeed, neoclassical economics' positive approach, justified by the two authors' cultural background. Similarly, new finance theory (hereafter, neoclassical finance theory, NFT) purports to develop positive models, but as a basis for normative ones (Jensen and Smith 1984).

The second feature of the paper is methodological individualism, consisting in the focus on individual behavior and the related reductionism (i.e. explaining by individual behaviors all social realities, including capital markets as well as organizations and institutions).[2] The predominant role of the single agent, with respect to the whole firm, becomes evident when one considers the paper's basic argument, i.e. the no-arbitrage principle. This principle claims that capital structure decisions of the firm are irrelevant since every single individual (investor) can reproduce by himself the firm's capital structure. Such a "home-made leverage" argument makes irrelevant (from the investor's viewpoint) every decision on how to finance the firm.

Methodological individualism is extensively employed by NFT. Here, methodological individualism comes to overlap the firm and the entrepreneur-shareholder. The main result of the application of this methodology is the development of a new theory of the firm which has no organizational content, since the firm is considered as an investment performed by single individuals, i.e. its shareholders.

The broader effect of methodological individualism has therefore been that processes taking place within the firm do not really exist at the organizational level, but only at the individual one. They are indeed individual decision making processes which must be consistent with – and checked against – individual decision making processes taking place on capital markets. This implies that the firm *as such* does not exist in NFT. NFT, like neoclassical economics (Machlup 1967), is focused on the price system. The most surprising effect of methodological individualism is then that the firm disappears from NFT. Even if the Miller and Modigliani paper is about the capital structure of a firm, it actually does not deal with any concepts of a "real" firm.

The third novelty is the way assumptions are formulated and employed. The irrelevance of capital structure and the "home-made leverage" argument hold under the hypothesis of "perfect" capital markets, which implies a frame of unrealistic (or even anti-realistic?) assumptions.

Drawing on Friedman (1953), the usefulness and validity of a theory have to be judged on the sole basis of its ability to make sound predictions. According to this idea, the assumptions founding a theory are not to be real but necessarily unreal. They are abstracted from reality for a predictive purpose and therefore they cannot be real. To be useful, an assumption must be false as a description of the reality. For example, assuming that the entrepreneur is a profit-maximizer does not mean that she operates in this way, but only that her behavior can be analyzed "as if" she acted according to the maximization principle (see Anthony and Biondi, this volume, for further criticism).

These three novelties are of interest for an historical analysis of finance theory, since they testify to a departure from traditional finance theory. The new theorization was born from the rib of the neoclassical economics, and therefore it inherited its methodology, especially the unrealism of assumptions.[3]

The introduction of this new way to do research was not without effects on academic circles. The American Finance Association (hereafter, AFA), founded in the late 1940s by a small group of economists from the American Economic Association (AEA), founded its own journal in 1946: the *Journal of Finance.* For some time the two spirits of finance theory, the traditional and the brand new, shared that journal. In 1967, the *Journal of Finance* edited a volume with the opening article authored by Weston (1967: 539), president of the AFA, whose incipit was:

> There is increasing evidence of division of opinion among members of the American Finance Association concerning the content and emphasis of financial research and teaching. Some may see in these disputes evidence of a disturbing lack of harmony in the Association. But they should also be recognized as a healthy sign of progress.

However, Weston continued emphasizing that analytical tools (mathematics, statistics, operating research, and so on) developed during and after World War II had begun to be applied to finance and other social sciences, and this by disregarding the advances in financial knowledge made by the traditional approach.

> The realistic and inexorable fact is that the content of the finance field has changed in the last decade. ... The emerging problems and issues of finance make it unsatisfactory for us to expect that we can contribute to the improvement of economic and business decisions solely by generalization and judgment. ... The older methodologies are useful for suggesting hypotheses and propositions, but inadequate for the systematic formulation of models and their testing.
>
> (Weston 1967: 540)

In the same volume, Sauvain (1967: 541), vice-president of AFA, replied:

> The emergence of the mathematical analysts and their confident invasion of the field of traditional finance present an even more serious problem among academicians than that of editorial policy of the *Journal.* The problem is: How do we live with them?

The hard question was answered some years later. In the 1972 the traditionalists founded the Financial Management Association (FMA) and a new journal: *Financial Management.*

NFT scholars, in their turn, founded in 1974 the *Journal of Financial Economics*, with the aim of publishing research articles that draw upon the financial

economics methodology. Despite the relevant development of financial economics – or maybe just because of it – a growing discontent began to arise. For example, Friend (1973: 257–58) in his presidential address to the AFA meeting, argued:

> Frequently, we specify assumptions which are known to grossly violate the real world facts . . . and justify our actions on the need to simplify theory and to appraise it only on the basis of its ultimate usefulness. . . . the trend in present research both in finance and in many other branches of economics seems to be more concerned with, or at least more likely to lead to, advances in methodological niceties than in substantive knowledge.

Later, he clarified that "by substance I mean solutions to the real world and not artificial problems." (Friend 1973: 272).

Friend seems to address the fundamental issue of the new research stream: the problem with the assumptions' unrealism and its consequences for the underlying understanding and view. The new way to do research in finance was and still is eminently based upon an axiomatic approach. Axioms are self-evident truths upon which theoretical propositions are built. Even if some assumptions seem to be acceptable when compared to the intended reality (as for example individual risk aversion), some others are clearly unrealistic. The issue noticed by Friend is then the main criticism against the new finance theory.

Notwithstanding the arising criticisms against NFT, much of the theoretical work in finance until the 1980s is based upon the "neoclassical" framework developed by economics and thus imported in finance theory, with some assumptions afterwards relaxed. The core propositions of such framework are the following:

1. economic agents are formally rational;
2. capital markets are perfectly competitive;
3. information is freely available and reflected by prices (capital markets are efficient);
4. the firm object is to increase or maximize shareholder wealth (the shareholders wealth maximization hypothesis);
5. and, with the development of the NFT, there are no-arbitrage opportunities.

Investors' rationality (point 1) implies that any agent seeks to maximize his own utility, which is formally described through a utility function based upon his wealth. To be tractable within the NFT models, such a utility function is assumed to be continuous and differentiable to any degree. In addition, investor behavior is assumed to be consistent with von Neumann and Morgenstern's (1944) axioms of rational choice. Individuals are able to identify all the possible outcomes of various decision alternatives, to evaluate them by assigning a certain utility degree, to identify the probability of each alternative, and finally to choose the one maximizing the expected utility. Such an investor behaves then as *homo economicus*, she is a synthesis of pure rationality and calculus.[4]

Many researchers have criticized the full rationality assumption because of its manifest falsity.[5] However it has been (and still remains) at the core of NFT on the basis of Friedmanian conventionalism. The investor is not required to be necessary an algorithmic function. It suffices that his behavior can be analyzed "as if" he or she is fully rational.[6]

The perfect market competitiveness (point 2) usually is the starting point of modeling. Markets are the benchmark for valuing the effectiveness of any agent behavior. Market competitiveness works in order to eliminate firms and investors not fully rational. Market forces impede the chances for noise traders to survive under the competitive pressure.[7]

The third core assumption is that (public) information is instantaneously incorporated into prices. In other words, prices reflect all available information and therefore it is not possible to systematically beat the market. Furthermore the price system acts as a coordination system of economic behaviors, since it provides agents with all the information they need for sound decision making.

The forth assumption reshapes, in different ways, the idea of profit maximization driven by neoclassical economics. Here the profit is spread along an inter-temporal period and therefore is translated in terms of discounted cash-flow value, which under perfect information translates into stock value. This leads to the hypothesis of shareholders wealth maximization (SWM), which merits a digression about the governance of the firm. The implicit NFT assumption before the Jensen and Meckling (1976) paper was that managers act on behalf of shareholders, and their decisions are straightforwardly oriented towards shareholder wealth creation. The Jensen and Meckling (1976) paper signals an important milestone in the development of finance theory, since it is probably the first paper to relax this notion of manager-shareholder harmony. The purpose of the paper is "to develop a theory of the ownership structure of the firm" (Jensen and Meckling 1976: 305) taking into account featuring issues of agency and property rights. One of the major findings of the paper is that capital structure is more than a simple way to allocate fixed cash flows produced by the firm. It is a very corporate governance mechanism, since it allows rights and duties among the organization members. As Brennan (1995) notes, a point often neglected in criticizing Miller and Modigliani (1958) is the *ceteris paribus* assumption regarding the income stream to be allocated. Since the income stream is assumed to be predetermined (exogenous), the irrelevance of the capital structure follows.[8] With the Jensen and Meckling article (1976) the theory of the firm enters overbearingly the finance theory. The firm then is seen a nexus of contracts, and this perspective is still at the core of finance theory (Zingales 2000).

Last but not least, another important milestone in NFT is the development of models based upon the no-arbitrage assumption (point 5). The no-arbitrage argument was already used by Miller and Modigliani (1958) in order to justify the notion of home-made leverage. In fact, the dissemination of that concept began around the 1970s with seminal articles by Black and Scholes (1972; 1973) and Merton (1973) on option pricing, and further developments by Ross (1978) and Dybvig and Ross (1987). The no-arbitrage assumption and the related the-

oretical models are of paramount importance for the development of the finance theory. As noted by Ross (2002), the very pillars of the "new" neoclassical finance are the efficient market hypothesis, the no-arbitrage assumption and the risk-neutral pricing. The founding of theoretical models upon the no-arbitrage assumption has been very useful for relaxing some basic assumptions of the "old" neoclassical approach, such as the form of rationality needed for operating in competitive markets. In the "old" neoclassical approach *all* investors were to be (or to behave as it were) fully rational in order to bring the market to equilibrium. According to the "new" neoclassical approach, only the existence of few fully rational agents is necessary, if each of them can take advantage of every arbitrage opportunity. As Ross (2002: 136) says:

> Neoclassical finance is a theory of sharks and not a theory of rational *homo economicus*, and that is the principal distinction between finance and traditional economics.

Furthermore, one of the major concerns of the "old" neoclassical approach, the market equilibrium assumption, is relaxed. Not only are some models based on local equilibrium, such as Black and Scholes (1972; 1973), but the general capital market equilibrium is now a possible result of the no-arbitrage approach. This way, states Ross (2002: 136) "we have developed the Newtonian version of our science."

## Lost in finance: finance theory and the firm

Since NFT is based upon methodological individualism, firm does not exist. There is nothing of what we know as firm in NFT. But why does firm not exist in NFT? And what is the firm we are looking for? Or to put it differently, what is the substitute for firm which lies at the core of NFT?

### *Firm and shareholders in the Miller and Modigliani proposition*

Firm in the Miller and Modigliani proposition is nothing more than a production function. As Miller notes (1986), firm for NFT is essentially the same as that for neoclassical economics, the unique difference being the fact that the latter is a production-function transforming inputs into outputs, while the former transforms today's money into tomorrow's. Then, "the two models of the firm, the finance model and the price theory model, are variations on a single theme" (Miller 1986: 452).

In the first stage of financial economics, therefore, there is no difference between the theory of the firm adopted by NFT and that in which neoclassical economic theory is rooted.

The firm is a simple production function, and it is held by a single entrepreneur or owner-manager whose main goal is profit maximization or, better, its expected wealth maximization. Under this assumption, NFT, before Jensen and

Meckling (1976), assumed that a separation between managers and shareholders is without effect on the shareholder wealth maximization (SWM) objective. Even if a manager could be thought of as self-interested, personal interests are of a second order with respect to those of shareholders. So, generally, there is no explicit reference to the problems arising from separation between ownership and control.

The basic decision making approach is firmly related to capital markets. Until the 1950s, the standard capital budgeting did not take into account any feedback from capital markets (Findlay and Williams 1980). No model of asset pricing existed. Starting from Miller and Modigliani's contribution and with the development of the Capital Asset Pricing Model (CAPM) (Sharpe 1964; Lintner 1965; Mossin 1966), capital markets entered the capital budgeting process through the price system. The relationship between investment projects and capital markets was established via the opportunity cost of capital. At the beginning, only capital markets were assumed to be in equilibrium, but developments of CAPM were able to encompass all the markets, real and financial, and the famous Fama and Miller (1972) book developed a finance theory where both real and capital markets are in equilibrium.

Since capital markets are the kingdom of single rational agents, individual behaviors become the reference for the decision making of any firm. Because of methodological individualism, individuals play the most important role within such a theoretical framework.

Finally, methodological individualism and equilibrium put the price system at the core of the finance theory. Any decision can be governed by continuously referring to prices. And the firm disappears from NFT.

## *Firm, shareholder, and agency theory*

The Jensen and Meckling (1976) article marks the beginning of a new phase for NFT. Drawing on both agency theory and property rights theory, they provide a theory of the firm's ownership structure. As they make clear, the ownership structure concept is quite different from capital structure, since it does not refer only to the way the firm is financed but also to the share of ownership claims held by the entrepreneur-manager. Of course, this effort provides an approach to finance theory rooted in a theory of firm which differs from the pure neoclassical one. Jensen and Meckling (1976: 308) write:

> While the literature of economics is replete with references to the "theory of the firm," the material generally subsumed under that heading is not actually a theory of the firm but rather a theory of markets in which firms are important actors.

Accordingly, this new perspective can have a disruptive effect on finance theory. First, capital structure is no longer a way to allocate an existing and given cash flow stream, as in Miller and Modigliani's paper (1958), but it can affect the

income that has to be distributed. Because of the separation between ownership and management, and since managers are self-interested, there is no reason to believe that managers' behavior will always be directed to the SWM.[9] Here the problem becomes, how can incentive schemes be designed in order to reduce agency costs? Or to say it more easily, how can managers' behavior be aligned with the SWM objective?

Second, this agency problem exists not only between managers and shareholders, but also between different kinds of claimholders. In this context, debt can also be seen as an organizational control tool. Indeed, managers of firms producing large free cash flows could be tempted to invest them not to generate positive Net Present Value (NPV), but to obtain personal benefits. In such a context, debt is seen as a way to drain cash flows from the firm and thus prevent managers from wasting them.

From this perspective, it is not really possible to assume that expected cash flow is independent from ownership and capital structure, contrary to the Miller and Modigliani approach. Capital structure decisions become corporate governance decisions. In fact, with the Jensen and Meckling approach, the firm becomes a nexus of contracts, a mere legal fiction. Jensen and Meckling (1976: 310) emphasize the contractual nature of the firm:

> Contractual relations are the essence of the firm, not only with employees but with suppliers, customers, creditors, etc. The problem of agency costs and monitoring exists for all of these contracts, independent of whether there is joint production in their sense; i.e. joint production can explain only a small fraction of the behavior of individuals associated with the firm.

In Fama's (1980: 289) words,

> The attractive concept of the entrepreneur is also laid to rest, at least for the purpose of the large modern corporation . . . management and risk bearing are as naturally separate factors within the set of contracts called a firm.

## Back to reality: some critical remarks on neoclassical finance theory

The unrealistic assumptions on which NFT is based has stimulated many criticisms. This section will discuss some of them, related to: (1) The positive NFT versus normative theory; (2) the pretension of founding a finance theory value-free; (3) the basic assumptions of *homo economicus* and capital markets' efficiency; and, finally, (4) the missing firm.

### *Positive vs. normative theory*

As already noted, at its inception, financial economics was essentially devoted to building positive models. The major concern was not to prescribe models for

decision making or behavior (as traditional finance did), but to build positive models to make correct forecasts. With this purpose in mind, finance researchers, as well as their economics colleagues, began to employ the "as if" approach in order to build models capable of prediction.

But the sharp distinction between normative and positive approaches is utopian. As Jensen and Smith (1984) pointed out, in order to prescribe decision models and behaviors, one should be conscious of their predictable consequences. Positive theories are useful not only by themselves but also because they represent the basis for any normative theory. Purposeful decisions cannot be made without sound positive theories. The research program begun in the 1950s was then also directed towards this end.

Nevertheless, the relationship between positive and normative theories is more subtle and implicit. In the social sciences, it is not possible to accept the strong separation between object and observer as natural scientists do. Any effort at understanding the social reality, whatever it is, has an impact on the reality itself. Ultimately, the positive approach cannot exist at all. For example, the well known market anomalies studied during the 1980s seem to be reduced. Of course, this is the effect of analyzing those anomalies.

### *NFT and values*

The NFT approach pretends to be value-neutral, i.e. not to be influenced by any particular ethical view of the world – since it does not ask questions such as "what is good" and "what is bad." According to Friedman (1953: 4), "positive economics is in principle independent of any particular ethical position or normative judgment."

The role of values and beliefs in finance theory can be viewed from different perspectives. It should be noted that the choice of a specific methodological approach is a judgment itself, as it derives from an evaluation of various methodological options. As Frankfurter and McGoun (1999) point out, the finance methodology is totally quantitative because the researchers believe that such a methodological structure can capture the objectivity of the world, which exists (or is supposed to exist) independently from the observers (i.e. the researchers themselves).

This statement deals with the problem of the value-free science in a very direct way. NFT is intentionally positive. It employs statistics for understanding the world, and in this effort it looks after an aseptic data analysis. However, data as such are not enough and the researcher's judgment in interpreting them is always required. During the process of gathering, analyzing and interpreting data, values and beliefs play their role in many subtle and sometimes implicit ways. First, building a testable model means selecting some relationships that are (or are believed to be) more important than others. Rejecting the other relations is a way for introducing values and beliefs in the finance theory. Second, the selection of the significant level for accepting the null hypothesis is another example of how values and beliefs enter financial models. As Zeckhauser (1986: 438) said, referring to the dispute between behavioral finance and NFT:

> Elegant abstract formulations will be developed by both sides, frequently addressing the same points, but because there are sufficient degrees of freedom when creating a model, they will come to quite different conclusions.

The existence of implicit beliefs is also testified by the core terms used in NFT: efficient capital markets, random walk, fair values, and competitive markets. All of them are value-laden (Frankfurter and McGoun 1999). For example, the random walk concept means that no investor can systematically beat the market, since the latter behaves unpredictably. In this way, no differences between investors exist, all of them being equal in the sight of market. "Fair value" has the same equalizing effect. Fair value results from a transaction between two parties at arm's length, neither being privileged either by available information or by bargaining power.[10]

## *NFT and* homo economicus

Neoclassical finance theory's methodological approach has been criticized mainly for assuming the fully rational *homo economicus*. Criticisms have arisen from many sides:

1 studies on bounded rationality
2 behavioral finance
3 studies carried on by the Santa Fe Institute on agents' behavior in complex environments.

All these three patterns of research are concerned with fully rational individuals' behavior. Individuals are endowed with bounded rationality (Simon 1982; 1983) which renders impossible the behavior predicted by economics.

Many authors have challenged the propositions – which are at the core of economics and finance – derived by the full rationality assumption. Since Ellsberg (1961) and Allais (1953), the main concern has been to check the validity of that assumption upon which economics (and NFT) is based. These scholars found, among other things, that: (i) individuals are affected by cognitive biases,[11] such as over-optimism and the framing effect; (ii) markets are not so efficient as predicted (hoped?) by NFT (Shiller 2000; Thaler 1993; 2000); (iii) competitive markets do not necessarily select away noise traders (De Long *et al.* 1990; 1991a; De Long *et al.* 1991b; Shleifer and Summers 1990; Cutler *et al.* 1989; 1990); and (iv) individuals are not always risk-averse (Kahneman and Tversky 1979) since they favor risk-seeking behaviors when actual results are lower than targets or than their reference point.

Finally, the path of research inaugurated by Brian Arthur (1990; 1995; 1999) and his colleagues at the Santa Fe Institute in Pasadena[12] developed models where individual expectations are dependent of the expectations they believe other individuals are making. In this framework of reference and analysis, the behavior assumed as "normal" by economics only exists under some specific conditions.

***NFT and the missing firm***

The three paths of research mentioned above have at least two common points, namely: (1) the attempt to criticize the assumptions of economics and NFT, and (2) the focus on the individual as the basic unit of analysis.

Another important criticism can be addressed to NFT following the surprise of Simon's Martian looking at the Earth's surface with a special telescope which identifies firms as green areas and market transactions as red. After understanding that the greater part of the surface it can watch through its special telescope is green, "it might be surprised to hear of the structure called a market economy. 'Wouldn't "organisational economy" be the more appropriate term?' it might ask" (Simon 1991: 28, reprinted in this volume).

This latter criticism is the main theme of this chapter, and will be pursued in the rest of this section and in the following.[13] Our major concern is the fact that a firm must continue to perform for all stakeholders that interact with the firm and through the firm. Such a perspective differs from the contractarian view of the firm based on explicit contracts (agency theory) for two reasons at least. First of all, the firm is not simply a nexus of contract but it is an entity, a whole, a dynamic system which cannot be understood as the mere aggregation of its constituencies. Second, the liquidation of the firm does not simply imply the need for contractual renegotiation. It also determines the disappearance of unique qualities belonging to the firm itself, and which can be appreciated only if the entity nature of the firm is recognized.

The introduction of the entity-firm concept calls for a different perspective in finance. In order to appreciate such a new perspective, the following topics will be examined:

1 The NFT static approach and the dynamic approach called for by the entity-firm perspective.
2 The entity-firm as a system, a concept that addresses both the dynamics and complexity generated by the numerous parties constituting it, and the impossibility of value-additivity tenet – which is at the core of firm valuation according to NFT. The latter must be superseded by the super-additivity of the entity-firm as a dynamic system.
3 The shift in finance focus when the entity-firm concept is introduced and replaces the previous shareholder focus.
4 The bounded rationality which characterizes real human beings, and therefore shapes organizational processes.
5 The role of costs and revenues in identifying the entity-firm's incomes and results in opposition to values.
6 The distinction (or opposition, according to Anthony 1960, reprinted in this volume) between the goals of firm continuity and the SWM.
7 The way to deal with uncertainty and risk and the departure from the concept of systematic risk as linked to the individual agent's behavior.

8 The consideration of law as a legal framework and a number of permitted institutional arrangements, within which the firm operates.
9 The role of cost as indicator of financial needs that firms have to cover the running of their business, in opposition to the role played by values within the capital structure decisions according to NFT.

### *Statics versus dynamics*

Even though the goal of SWM in place of profit has apparently brought time in the analysis, NFT still remains essentially static.[14] In fact, NFT deals with value as the instantaneous weighted mean of expected future cash flows (Biondi 2005b). Therefore, the future is already inscribed in present values (and current prices), and time – just introduced into analysis – is suddenly frozen. On the contrary, dynamics does not call only for time *per se*. It calls instead for the interaction between historical and future patterns that develop through time and create path-dependent phenomena. Then, an historical approach is necessary to manage firms, where "historical" means to understand the present and to manage the future while being conscious of the complex relations of the firm with the past.

The essence of dynamics lies in investigating the diachronic streams of costs and revenues, whose matching is required for satisfying all undertakers period by period (see Biondi 2005a for further developments).

### *System, complexity and value-additivity*

Entity-firm has a systemic nature, in the sense that it exists and can be understood as a system of relationships between all resources – material or personal, financial or economic, tangible or intangible – provided by all the actors undertaking the business.

Since the firm exists as such only through relationships, it is not possible to look at the firm by simply looking at each single resource which is a part of it. Within the firm, single and independent resources do not exist since they are interrelated. Such a conception leads to a different view of the firm. It is no longer a mere aggregation of assets or resources. As a matter of fact it is a complex system of these and of their mutual relations. Relationships among resources, when considered in the same instant and also over time, confer on the whole firm a different economic meaning than that attached to the single resources separately. NFT instead focuses on each single resource, either tangible or intangible, and therefore ignores the difference between a resource considered in isolation or as a part of the entity-firm system.

The entity-firm concept appears to be useful especially nowadays, when the competitive position deriving from single resources possession and control (not only from their property) is less important than that deriving from the way resources are combined in a systemic way.

### *From shareholders to the firm*

According to NFT, any behavior should be shaped by the shareholders' wealth maximization. On the contrary, under the entity-firm approach, drawing on the satisficing principle fostered by Simon and Anthony, management's primary goal is to guarantee the firm's continuity whilst producing at the same time satisficing compensations (or returns) for resources undertakers have provided to the entity. This satisficing principle should not be confused with notions referring to compensations determined under an equilibrium framework. "Satisficing" means that each compensation (or reward) must be harmonized not only with any other compensation given to every other stakeholder, but primarily with the necessity to safeguard the ongoing functioning of the firm. "Satisficing" further relates to a bounded rationality context. Finally, it is not possible to identify such a thing as the optimal or equilibrium compensation. Compensations refer to a satisficing degree which appears to be adequate at the moment the relation is taken into consideration.

### *Bounded rationality vs. full rationality and the homogeneity of expectations*

Under NFT, investors are (or, in a Friedmanian perspective, are thought to behave "as if" they were) fully rational. Moreover, they share the same expectations about the future. Each individual is then a standard algorithm without emotions and passions,[15] and enters the model as a standard processor which gathers and elaborates information. There is no room for expectations un-homogeneity in this context. The value of every asset is the same for every individual, and it conveys all information about its future usage value. This means that, under NFT, the value of every asset is defined according to the maximal revenue that a unit of that resource would make possible in any conceivable usage (Denrell, Fang and Winter 2003). Thus, once such a value translates into the equilibrium price through efficient markets, it conveys to investors all the information about how much the best usage is worth, so making it possible to evaluate investments, and take perfect decisions.

In contrast to NFT, the entity-firm approach is based on uniqueness, i.e. on expectations un-homogeneity and diversity. This is not a self-serving approach, but one which is grounded on the realism of the assumptions. In this way, the difference between values (i.e. the sum of discounted expected cash flows, according to the definition provided by NFT) and both costs and revenues (as defined by accounting) assumes great importance. The latter, indeed, are the results of transactions where bounded rationality and uniqueness play a key role.

### *Costs, revenues, and values*

The difference between finance-values and accounting concepts should not be understood as the result of market imperfection or inefficiency. This would be

correct only if reality were grasped by the unrealistic lenses of NFT. That difference, instead, is implied by the uniqueness of the firm and bounded rationality together, not only by accident or different expectations (Barney 1986). Since each firm has its own uniqueness, the value of any asset becomes specific to the firm itself. The firm's incomes result from the joint application of that asset's specificity through business combinations and transactions, not from a universal asset value which is the same for everybody.

Costs and revenues are at odds with the NFT concept of values, both in their ontological and in their epistemic nature. From the ontological standpoint, they are determined within a bounded rationality framework. Such boundaries manifest themselves both: (i) as the impossibility of defining the "true" value for each resource depending on its actual and potential use; and, (ii) in a relatively situated contest, as the impossibility of negotiating the best value (i.e. the minimum cost or the maximum revenue) because of the partial knowledge about the parties potentially interested in undertaking the transaction. Therefore, the difference between accounting concepts and NFT value is quite marked. According to NFT, value is universal because of rational individuals and perfectly efficient markets, and it is therefore the fairest and, at the same time, the best. In real life, many different costs and revenues potentially exist at the same time and over different periods.

From the epistemic standpoint, the matching of costs and revenues, as determined through the actual transactions accomplished, identify an income unique for each firm. In NFT, the value of any asset being universal, any fictitious firm is expected to earn the right (fair) return with respect to its risk. According to NFT, positive economic profits cannot stably exist since they are at the odds with the underlying assumptions. Under the entity-firm approach, the firm can obtain net earnings which are very large – larger than any supposed equilibrium return – and the concept of "economic profit" as lies at the core of NFT and economics cannot exist, since it is at the odds with entity theory's basic premise: the uniqueness of the entity-firm.

The entity-firm approach does not deny any meaning to value-based concepts. In fact, it relates values to the underlying streams, costs being a type of stream, revenues another. Then, it is the continuous comparison between costs and revenues taking place at any time that highlights the firm's uniqueness. The difference between cost-stream and revenue-stream testifies to the firm's uniqueness in the eyes of its stakeholders.

### *Firm and shareholder wealth maximization*

In NFT the firm has a sole objective: to maximize shareholders' wealth. Shareholder wealth maximization (SWM) translates the profit maximization principle into operational terms. In the pure competitive market, indeed, the firm is nothing else than a profit-maximization reactor (Machlup 1967). As pointed out by Solomon (1963) and Weston (1966), profit maximization is not useful as an

operating criterion. Profit maximization is vague, does not consider intertemporal comparison of profit streams, nor the quality (riskiness) of them.[16]

What is the rationale for such an objective? From a narrow perspective, it could be argued that the shareholder is the firm's residual claimant, and since it is not possible for him to write a complete contract with all parties, his claim and related supremacy are the other side of the risk he must bear (see further criticism by Stauss (1944), reprinted in this volume).

Even NFT's proponents highlight of course that maximizing firm or shareholder value is an instrumental objective, the ultimate one being the enhancement of social welfare.[17] Using Jensen's (2001: 302) words:

> The real issue to be considered here is what firm behavior will result in the least social waste – or equivalently, what behavior will get the most out of society's limited resources – not whether one group is or should be more privileged than another.

Under the theoretical assumptions which NFT is based upon, there is no contrast between SWM and a fair compensation for all the undertakers or stakeholders. Under those assumptions, SWM is an instrumental objective for reaching the firm's real purpose: the maximization of social welfare (Copeland 1994; Jensen 2001). Under those assumptions, then, there is really no difficulty in accepting that goal. Nor, to justify SWM, is it required to argue that shareholder, consumer or worker are not specific subjects but just roles which are played by individuals (Dufrene and Wong 1996). Under the theoretical NFT framework, SWM is all that is needed to enhance social welfare.

But the validity of SWM relies on both perfect competition and full rationality, i.e. on a world where oligopolies and externalities do not exist – though they factually do. A world where individuals are (or behave "as if" they were) fully rational – though if they are not. Therefore, the major criticism of NFT does not apply to the supremacy it assigns to shareholders, which is nothing more than an instrumental objective, but to the assumption that SWM is the best that can be done to improve social welfare.

Leaving aside the abstract framework provided by NFT to enter a reality-driven perspective, the problem is really how to achieve social welfare in a world where markets are not perfect and individuals have only bounded rationality.

Since a firm is a means for all stakeholders to achieve satisfying results with respect to their needs, the shareholder supremacy problem must be revisited from different standpoints.

First, the equation "maximizing shareholder wealth = maximizing social welfare" can only be accepted when oligopolies and externalities do not exist, which is very far from reality, and therefore prevents one from accepting the equation.

Second, even if there were not oligopolies and externalities, or if in the future one could expect (or hope) this, maximization requires that individuals are endowed with full rationality. Since individuals are – and they are expected to

go on being – not fully rational, their behavior being affected, for instance, by cognitive biases (Kahneman *et al.* 1982), wealth maximization can at least be thought as an intended rather than an actual behavior. Since maximization is not really possible, one cannot expect that any decision, even if intended to be optimal, would be "as if" individuals were value-maximizers in the meaning assumed by NFT. Individuals can be at most intentionally maximizers, but this does not mean that every decision results in a maximizing behavior.

Third, SWM implies that individuals have a long-term evaluation of the firm. From this perspective, there could be a strong coherence between SWM and the entity-firm approach, which focuses on guaranteeing the firm continuity whilst providing fair returns and compensations for all stakeholders. In fact, SWM calls for an omniscient human being, but it is difficult to imagine that such a being could ever exist.[18] On the contrary, in order to guarantee the firm ongoing and satisficing returns for all stakeholders involved in it, specific attention to this point is required. Then, it is simply impossible to take for granted that simply looking after shareholders' wealth in the long term already implies the rest.

There is another reason for rejecting the apparent relationship between SWM and firm continuity. According to NFT, the relationship between firm and stockholder is limited to an investment decision where the only important dimension taken into analysis is the risk-return trade-off. From the NFT standpoint, each firm is the same as any other with the same trade-off, and some conceptual tool exists to appropriately compare each firm with another. For example, CAPM-Beta can be used to "measure" each firm in terms of each other. From the entity-firm perspective, however, relationships that shareholders cultivate with and within the firm are usually not replicable with or within another firm. Of course, various categories of shareholders can be identified, ranging from NFT-like shareholders – whose interest in a specific firm is very low since they are interested uniquely in the risk-return trade-off – to shareholders whose interest in the firm is very specific. The more the implication is relevant, i.e. the more a stockholder is involved in a specific firm, the less can he switch or compare his position to another firm with the same beta.

### *Value, risk and equilibrium*

The fundamental and well known tenet of NFT is that only systematic risk calls for compensation. Such a proposition is the result of the famous CAPM models developed independently by Sharpe (1964), Lintner (1965) and Mossin (1966), and it is based on the diversification principle on which Markowitz (1952) built portfolio theory. The CAPM divides the overall risk of any security into two parts. Unsystematic risk usually refers specifically to a firm, and can be canceled out through diversification. Systematic risk depends on the economic system as a whole and cannot be diversified. In the CAPM systematic risk is linked to expected return on the security through the beta of the security itself.

Since unsystematic risk can be canceled out, investors are not affected by it. When valuing securities, investors only take into account the systematic risk that

actually affects expected payoffs. Managers should not be concerned with unsystematic risk. Such a part of risk can be in fact eliminated by the investor-shareholder herself and is not a topic requiring attention by managers. Application of CAPM for valuing investments then translates into such an imperative for managers to consider only systematic risk.

However, the application of CAPM-like models to the capital budgeting process can be criticized.

First of all, the existence of equilibrium cannot explain the existence of positive NPV investment. If markets are in equilibrium, the expected return on any security is exactly appropriate to its systematic risk. In such a case, the return is the right compensation for bearing the (systematic) risk: no positive NPV projects may exist. Positive NPV means that a project is expected to produce a return greater than the equilibrium level, which is determined according to its risk. But this means that the risk-adjusted rate of return employed for valuing the project is not the correct rate, since in this case NPV would be equal to zero. Accepting the existence of positive-NPV projects does not permit us to clarify if we are in a situation of inter-o intra equilibrium (Findlay and Williams 1980).

Another interesting criticism has been addressed by Bettis (1983), which identifies the conundrum of strategic management when combined with the systematic risk. The author highlights that while NFT preaches the importance of the sole systematic risk, strategic management is interested in the firm's overall risk management. Such a remark is of interest for the purpose of this chapter, not in order to compare strategic management to NFT, but because it pays attention to the reasons supporting overall risk management.

First, even if it was possible to assume (and to consider it acceptable) that shareholders are fully diversified[19] so that only systematic risk matters, other agents, such as managers and workers, are instead poorly diversified. Their principal investment is represented by the structured relations they have with the firm and which take the form, for example, of labor relations. The flows of investment related to these relations are usually linked to one single firm. Even though not-specific shareholders are fully diversified, workers are not and, from this perspective, they need to manage that part of the overall risk that, being unimportant for most shareholders, is from their viewpoint of paramount importance.

Another way to cast the same doubt is to consider that the firm is not a mere legal fiction but an entity having institutional content, i.e. a dynamic system of somewhat stable relationships inside and outside its legal frontiers, whose primary goal is to continue indefinitely whilst satisfying the multiple and various interests it conveys. This perspective opens to further analysis the importance of unsystematic risk. If market competitiveness could be seen as a mechanism for improving the economic system's efficiency, from the viewpoint of any single firm it is quite similar to the sword of Damocles. If, from the diversified shareholder viewpoint, the failure of a firm is balanced by the total return on his portfolio, from the firm's viewpoint it is its downfall that matters. Then, whenever

looking at the firm not as a simple veil but as an ongoing system, the management of systematic risk alone cannot be accepted.

This point may be further developed by recalling the factors from which unsystematic risk derives. One of them is technological innovation. Let consider an R&D project valuation. Following NFT, it should only consider systematic risk for valuing the project. A firm that renounced the management of unsystematic risk, technological evolution, and in general any process of economic innovation, could exist only by chance. This would be of null effect on the shareholder, since he diversifies, but could be dramatic for firm-entity continuity, especially in a competitive and evolving context.

Summarizing, the way NFT deals with risk is therefore partial, in the twofold sense that it is based on shareholder supremacy, and it is limited. Shareholder supremacy, as said, translates into neglecting the firm as an entity. A firm is simply a type of investment, like a rent, and therefore the shareholder is only interested in a compensated risk-return arrangement. Since NFT does not recognize the firm as existing separately from its shareholders, all firms can be understood by simply comparing the expected return together with the expected risk. Therefore, the fundamental question about how to deal with risk knows a sole answer: the diversification principle. Such a principle shows precisely the shareholder's supremacy, but at the same time the limitations of NFT, since it does not address the role of the firm in coping with uncertainties and complexity.

From its limited perspective, NFT assumes that only systematic risk matters, since only such a risk affects the return required by shareholders to invest in stocks or other securities. Any single shareholder can protect himself from the total risk of his investment by a simple diversification process. Such a decision takes place on capital markets and is carried on through individual strategies, in the sense that each agent can efficiently and effectively diversify her portfolio. Therefore, NFT does not allow the possibility of managing risk at the firm level, through the operational processes which take place within the firm. This means, for example, that the strategic diversification processes run by firms are assumed to be either unnecessary – since the same results could be obtained directly by single shareholders, or wasteful – since those processes are assumed to be driven by managers only out of self-interest.

On the contrary, once an entity-firm approach is followed, those processes may be used as a way to protect shareholders (as well as all the undertakers) against the risks the firm deals with.

### ***Firm, finance theory and the law***

The firm in NFT is set out of law. NFT usually does not consider such strange things as law and institutional framework. But law influences many decisions related to the firm. The simple case of dividends can clarify the point. NFT assumes that free cash flows (i.e. cash flows residual after all investments in positive-NPV projects) must be distributed to shareholders. Here it does not matter why the shareholder is the privileged subject. What is

under investigation is the fact that all limits to dividends are usually ignored. In most countries, instead, law prescribes that dividends can be distributed to shareholders only up to the business income generated, as represented by accounting earnings, once a certain percentage of that has been allocated to reserves (Bline and Cullinan 1995; Biondi, this volume). One could argue that there are many ways to distribute free cash flows to shareholders, one of these being the share buy-back. But these kinds of operations are also subjected to legal constraints.

Moreover, there is a theoretical distinction between produced-realized earnings and distributable cash flows limited by established earnings, and such a distinction relies on the defense of the firm's continuity. Such a distinction is fruitful in order to put the ongoing firm at the core of financial decisions.

The following sections will draw upon and further develop this point.

## *Capital structure decisions, values and costs*

According to NFT, a focus on values together with the capital markets' efficiency assumption creates a strong overlapping between values, which are calculated on the basis of estimated/expected cash flows and are therefore only potential cash, and actual cash. To use accounting jargon, values are cash-equivalent. Throughout this chapter such equivalence will be referred to as the "liquidity principle."

This principle produces two main effects:

1 Since estimated values are nothing less than cash, any investment with a positive NPV will ever find "provisions"[20] for covering its financial needs, since capital markets recognize that such a positive cash-equivalent value can be easily translated into "real" cash.
2 The capital structure decisions are not centered on the concept of *cost invested*, as a quantifier of financial needs, but on the concept of value which is, as stated above, the driver of every decision.

Under NFT, financing a positive-NPV project is not difficult, since perfect and efficient markets really know that the project has a positive NPV and behave accordingly. The problem of project financing, therefore, is not about raising funds, since funds for positive NPV always exist in such capital markets. The problem is the debt/equity ratio choice. It should be clear in fact that, in such a framework, capital structure decisions do not deal with financing the firm in the sense one could imagine whenever referring to real firms. Those decisions deal with the firm's value distribution or with the governance system either of the firm and/or of its investment projects.

In NFT, it does not matter, for instance, if expected cash flow is not able, at some point, to cover the payments required for debt service. Since markets are perfect and individuals are fully rational, it is always possible to find someone willing to refinance those payments.

## Financing the entity-firm

Since the entity-firm's goal is to continue indefinitely while generating satisficing incomes, the financing problem can be approached by analyzing how finance can support the firm's continuity while incomes are created for all undertakers. Financing is therefore a second-order problem, in the sense that the primary object of decision making is not capital structure *as such*, but the firm's ability to manage the overlapping dynamics and matching between costs and revenues, which in turn constitute a complex and inextricable system. Financial needs depend on this costs-revenues dynamic correlation. Usually NFT does not explicitly recognize the temporal gap between costs and revenues which factually characterizes the firm's functioning. Such a gap is not the primary concern of NFT. Since the liquidity principle, NFT only evokes the financial gap, but never deals with it.

Once the firm's continuity is put at the core of business finance, some apparently counterintuitive and sometimes paradoxical situations can be understood. Hereafter, the following issues will be analyzed:

1 the role of financial slack in sustaining the firm's goal to continue, in opposition to the tenet of NFT which identifies the slacked resources as hostage of selfish managers, and therefore wasted with respect to the SWM goal;
2 the different meaning of capital structure decisions under the entity-firm approach, and its difference with the coalition approach;
3 the paramount role of costs in the financing process, which has been already sketched in the previous section;
4 the abandonment of any attempts to find a statically (*ex ante*) optimal capital structure, and a new focus on a dynamic satisficing capital structure;
5 the acknowledgement that capital structure decisions span over time, thus featuring dynamic substance.

### *Financial slacks*

According to NFT, fully rational human beings can chose *ex ante* the best course of action given their ability to forecast and comparatively valuate all the alternatives. Within such a framework, the probabilistic approach to uncertainty allows one to fairly define the certain-equivalent for any stochastic distribution of payoffs, so that any actual decision is coherent with the probability structure of uncertainty.

In fact, bounded rationality is connatural to human beings. Not only they are usually not able to manage the intricate interaction of numerous variables, but they feature a natural ignorance of the future. The possibility of translating known possible payoffs into fair certain-equivalent, if any, is therefore limited.

When faced with the uncertainty surrounding the future, financial resources slacks can be accumulated in order to preserve the entity-firm's continuity against the incorrectly computed equivalencies of uncertain and certain values. Provisional slacks of resources therefore may be interpreted in a different way

than NFT does. They permit the firm's continuity since they act as cushions against crises, which might occur because of both forecasting errors and the natural ignorance of managers and other undertakers.

Even though not every slack can be justified from that standpoint, they are not, however, all waste. But it is clear that once the firm's primary goal has to be SWM, all slacks should be considered wasted resources to the extent that they allow either (1) the independence of managers from shareholder control exercised by denying funds; (2) the sub-optimal (not maximal) return for the shareholders, who in turn ought to be left to take the best decision – for their personal benefit – about how to invest that money.

Under NFT, slacks are beneficial only for managers who can manage the firm according to their own interests. Thus, the problem becomes how to bring back such financial resources to shareholders. The free cash flow concept has been used by Jensen (1986) in order to discuss agency costs linked to the availability of those resources. Free cash flow is the cash flow in excess of that required to fund all projects that have positive NPV when discounted at the relevant cost of capital. Besides the fact that: (1) the free cash flow definition/calculation depends on how many and which projects have positive NPV, and (2) self-interested managers could identify as positive-NPV those projects that actually are not, such free cash flows should be distributed to stockholders in order to increase their expected wealth.

Under the entity-firm conceptualization, instead, shareholder wealth is not the firm's primary purpose. Shareholders can be at best *primi inter pares* (first among equals), but in any case the firm's continuity has to be privileged. From this perspective, resources slacks can be intentionally created to deal with unexpected larger cash outflows and/or smaller inflows experienced under real dynamics and complexity driven by bounded rationality and ignorance of the future. When the focus of business finance is the entity-firm, the necessity to finalize any action to the firm ongoing, comes to interpret such cash flows as not completely free, contrary to Jensen's views. The entity-firm's interest could require the creation of slacks in order to preserve the entity during financial crisis, even if that crisis is only expected and the slacks provisional.

NFT proponents could argue that such an approach cannot be justified, by resting on one of the two following assumptions: (1) positive-NPV projects always can find money on capital markets (the liquidity principle); and (2) the potential financial crisis can be predicted/forecasted in order to provide resources when they are needed.

Counter-arguments may be based mainly on bounded rationality and asymmetric expectations/information. Because of these factors, the liquidity principle does not always hold (point 1), nor is the forecasting ability so wide (point 2).[21] Therefore, such slacks become essential in order to deal with the anti-realism of both the liquidity principle and the full rationality assumption.

The argument that financial slacks are not always detrimental to the firm's value also relies on the importance of a entity-firm's autonomy *vis-à-vis* other firms (especially banks). From such a standpoint, the firm has to: (1) generate and

retain sufficient cash flows in order to undertake valuable investments; (2) produce adequate compensations for funds provided by external sources, in order to support stable relationships with their providers.

The second point seems very similar to the NFT tenet of a minimum rate of return required by investments. Such similarity is, however, only apparent, since in this case it is not assumed that such a rate is an equilibrium-opportunity rate.

The first point appears to be at the odds with NFT. That the firm has to generate and retain free cash flow is based upon the assumption discussed above about the need for financial slacks. But from the bounded rationality standpoint, it also relates to the need to save actual resources for the future. Such internal resources, that NFT views as a way managers have to increase their own wealth (e.g. perquisites) at shareholders' cost, are instead of paramount importance in supporting the independent ongoing functioning of the firm.

## Capital structure, entity-firm and coalition

From the entity-firm perspective, business' finance goal is to support the firm's continuity. This might appear at first glance to be a case of old wine in a new bottles, since the firm's continuity would be a way to focus implicitly on the wealth of single coalitions or members inside the firm. The difference is that the whole firm never coincides with some inside coalition, even if such a coalition can sometimes have more controlling power than another. What the theory of the entity-firm really stresses is the firm's continuity as such and not as a representative of some dominant coalition.

### *Financing the values or financing the costs?*

Whatever approach is employed, capital structure decisions in NFT are always based on the firm's expected value. What drives any decision about how to finance investment projects is expressed in terms of values deriving from discounting expected cash flows. Such an approach can be undoubtedly fruitful when capital markets are characterized by the "liquidity principle" of any project or firm (often seen as a collection of projects). "Liquidity" refers to the possibility that any investment decision can be expressed in terms of values, since such (potential) values can be transformed into money at any moment, by selling out or refinancing.

While NFT valuates capital structure decisions by considering their effects either on the cost of capital or on the firm value, within the entity-firm framework they are evaluated from the standpoint of financing the firm. This means that costs invested are at the center of financing since they quantify resources.

But in this effort, the funding of operations does not call for calculating their present or future values (this is a matter of what activities/projects have to be undertaken), but for providing provisions to cover costs. It should be noted that the cost concept stressed here is not the opportunity cost concept used by financial economics. Here, cost is the transaction-based consideration of the

asset acquired, and, even if it could coincide by chance with the opportunity cost determined within an equilibrium framework of prices–costs, it usually does not. Such a transaction-based cost concept is worthwhile because it allows the firm to settle inward exchanges of resources, ensuring that the firm's uniqueness prevails over the different values every firm attributes to the same resource.

Anyway, what has to be financed is the cost of purchasing the resources and not their value. In this way, not only is the firm appreciated for its uniqueness, but also such uniqueness is strictly related to the financing needs. Therefore, accounting systems can provide a meaningful story of the firm. Costs invested are taken into account through that system, and represent the primary source for recognizing the firm's financing needs. The accounting system can therefore provide a useful representation of the firm's financing behavior, and also a reliable basis for new decisions (see Biondi, this volume, expanding on these points from a dynamic accounting view).

## *A dynamic satisficing capital structure*

Under bounded rationality and firm uniqueness, it is not possible to fix an objective capital structure (as proposed by Modigliani-Miller and NFT) only by referring to value as an abstract concept, since such an approach does not consider the firm as unique. A satisficing financial structure can be found by relating cost-revenue dynamics to cash flow dynamics, as can be inferred by the strategic analysis of the firm as embedded in the economic system in which that firm operates.

Under the entity-firm theory, debt and equity (or to put it a better way, the different categories of debt and equity) are important for the different patterns of compensations they claim according to the law.

According to Jensen, one might suppose that firms operating in mature markets produce cash flows so relevant that debt can be used in order to reduce agency costs related to the abundant free cash flows generated. From the entity-firm perspective, instead, debt is a viable way of funding because of those vast cash flows produced or expected. From the standpoint of the conclusions, both approaches arrive at the same prescription (i.e. to substitute debt to equity), but justify it for different reasons. What makes the difference is not here the arrival point, but the underlying perspective. According to NFT, debt is a control mechanism against self-interested managers. Under the theory of entity-firm, increasing leverage is acceptable whenever it does not threat the firm's continuity. A stable firm's position within its sector can be translated into the opportunity to substitute debt for equity and thus leverage the return on equity, taking into account both the relationships between the costs and the level of debt and the possibility that a sudden fall in revenues and/or operating income producse negative effects on the firm's situation. Instead, according to NFT, that opportunity does not exist, since the increased risk provokes an offsetting increase in the rate of return required by investors.

Similarly, the case of a firm at its early stage of development may be explored. At that time, free cash flow is usually negative. From the entity-firm perspective,

the equity-based financing is compliant with the fact that firm does not have the potential to serve the debt. Then, equity is more effective in financing the developing firm, since it does not oblige the firm to make capital repayments for its service. From this perspective, the financing forms and their attributes are relevant according to the characteristics of both financial needs and cash flows generated by the ongoing firm. This comes to a perspective related to business strategy, which makes it possible to understand how financial needs and operating cash flows, provisions characteristics and legal requirements (which establish claims and obligations in servicing those provisions) are more suitable for financing the entity-firm.

### *Dynamically rebalancing capital structure through time*

To go further on capital structure decisions according to the financial theory of the entity-firm, let consider a developing firm, whose future financial needs are expected to behave as the mapping graph A in Figure 17.1. The three segments a, b and c show how – over the period and depending on expected financial needs – the firm should increase the share of permanent provisions with respect to the temporary ones. This means that the satisficing capital structure can be considered not only from the debt/equity ratio perspective, but also from that of the temporary/permanent ratio.

Figure 17.1 also represents the basis for some further insights. First, there is the progressive managing of the provisions mix. Until the time t1, permanent provisions level is set equal to segment a, while temporary provisions are equal to segment b. Once the firm functions and goes on, the mix between permanent and temporary provisions may and will change. Because of bounded rationality,

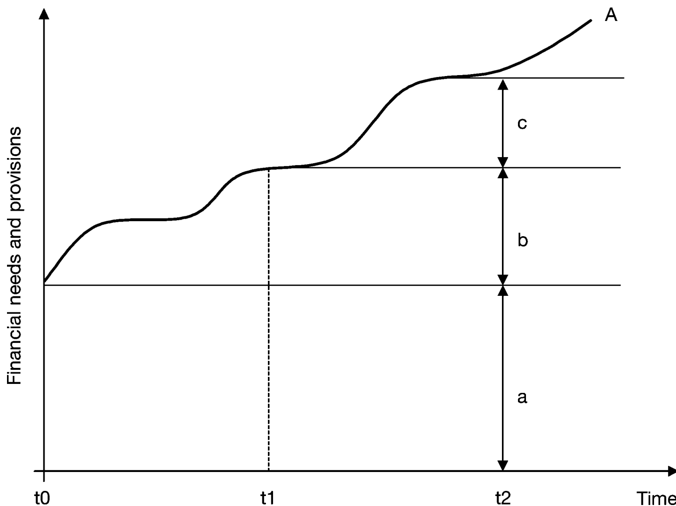


*Figure 17.1* Financial needs and provisions.

however, it is not possible to perfectly anticipate the quantitative and qualitative evolution of future needs.

Moreover, the ratio of permanent/temporary provisions cannot and is not instantaneously modified. The entity-firm operates dynamically and continuously, thus the ratio is modified over the time in relation to: (1) actual and past dynamics, and (2) business expectations concerned with the whole entity, expressed by management and by other constituencies.

In this framework, an optimal ratio does not exist as a universal level, since the ratio is continuously modified by dynamic evolution and complexity. Nor does it exist as a specific level, since it is not possible for the firm to identify any but satisficing ratios. In a world of bounded rationality and path-dependencies, it is actually impossible to enssure (and thus to *positively* assume) any optimal behavior.

The dynamics and complexity of the real entity-firm, with respect to the fictitious firm developed by NFT, can also be appreciated by highlighting that the distinction between temporary and permanent provisions is a mere simplification, since different types of provisions are characterized by various degrees of duration. The entity-firm has to choose not only between permanent and temporary sources but also between different degrees of duration. Under NFT, this choice is abstracted away and positively does not exist due to the assumed "liquidity principle." But for the entity-firm the problem is factually relevant.

Furthermore, the decision about the debt/equity ratio is not independent from the financial needs characteristics, contrary to the prescriptions advocated by NFT. In the real world (and from a realistic view), one can not fix the debt/equity ratio without reference to the characteristics of expected financial needs. This problem could at least be analyzed as an operating problem, but more probably it falls beyond the interest of NFT. After all, since the firm is constituted such that positive-NPV financial gaps will be always funded, the distinction between temporary and permanent provisions comes to be irrelevant.

Finally, because of bounded rationality and asymmetric expectations/information, the entity finance view claims that the business entity is confronted with such decisions, i.e. the choice between debt and equity and between temporary and permanent provisions. Figure 17.2 illustrates how the distinction between temporary and permanent provisions echoes that between debt and equity, even though there is not a complete match between temporary debt and permanent equity, since different types of debt and equity have different degrees of duration, more or less temporary or permanent.

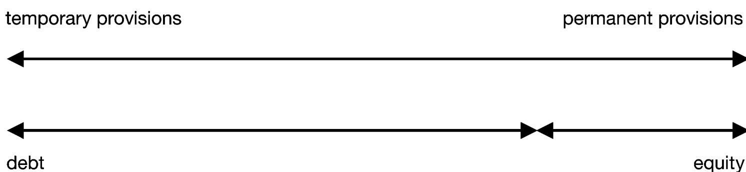


*Figure 17.2* Temporary and permanent provisions.

## Conclusions

This chapter has analyzed how the entity-firm approach calls for a dramatic change in business finance theory. In order to offer some insights about such a revolution, the chapter has addressed some fundamental criticisms of NFT, and focused its attention on methodological individualism which has abstracted away any idea of the firm as distinct from a mere nexus of contracts or shareholders' claims.

The advocated centrality of the entity-firm concept, instead, appeals for a business finance theory based upon preservation of the firm's continuity. Such a different keystone offers a different understanding of financial slacks, of values and risk, and finally of capital structure foundation and evolution. This latter topic has been developed throughout the chapter in a dynamic and complex context where bounded rationality and ignorance of the future impose a continuing rebalancing of debt/equity and temporary/permanent financial sources.

## Acknowledgment

I would like to thank Yuri Biondi (CNAM, Paris) for his valuable comments on an earlier draft of this chapter.

## Notes

1 In 1963 the same authors published a "correction," where they took into account the existence of corporate income taxes, whose effect is able to modify the irrelevance proposition. In 1977 Miller proposed a model of analysis where personal income taxes were also considered.
2 Allen (2001) notes that financial institutions are still ignored by finance theory.
3 It is noteworthy that the new neoclassical finance, mainly based on the no-arbitrage approach (Ross 2004) seems to be more parsimonious with these unrealistic assumptions, probably because they were and still are at the center of numerous criticisms. However, the elimination of unrealism from finance models is just taking its first step (if any).
4 For an early criticism see Veblen (1898).
5 See fourth section below.
6 With the development of the new neoclassical finance (Ross 2004) the assumption that all individuals are fully rational has been relaxed. What is needed is that only few rational investors exist, who can benefit from any arbitrage opportunity, and so bring the market to equilibrium.
7 See fourth section below for some critical remarks on such a proposition.
8 If income streams are invariant with respect to capital structure they are determined exogenously with respect to the variables which the model is built upon. This point is the basis of criticism developed by proponents of property rights theory, agency theory, and by neo-institutionalists. When contracts are incomplete, in fact, income is endogenous to the firm and depends on the governance structure. It should be noted that within the entity-firm approach proposed throughout this chapter, income is endogenous to the firm not because of its governance structure but because it is one of the constituents/fundamentals of the business entity.
9 On the same line of reasoning, Myers (1977) develops an analysis of the underinvestment problem.

10 An analysis of the values embedded in the finance theory can be made from a different perspective, that is, the effects that finance education has on students' perceptions, beliefs, attitudes, and behavior. Many studies (see Ardalan 2000a; 2000b; 2004 for a review) on this topic found that students trained in economics are more individually and self-interestedly oriented, less honest and more confident in the fairness of market results than their colleagues trained in other disciplines.

11 Here it should be noted that using the term "bias" to indicate some deviations from standard decision theory implicitly assumes that such a theory is substantially correct.

12 See Mouck (2000) for a synthesis of these studies.

13 For an early conceptualization of the entity-firm, see Zappa (1937; 1954) and Nicklisch.(1932). See also Canziani (1987) and Canziani and Rondo Brovetto (1992), for a historical and methodological analysis, and Zambon and Zan (2000).

14 On similar position see Shubik (1993), reprinted in this volume.

15 As Copeland and Weston (1988) state in their famous book on finance:

> there is much one can say about the theory of choice under uncertainty without, for example, understanding why a 70-year-old person is more or less risk averse than the same person at age 20, or why some people prefer meat, whereas others prefer vegetables.
>
> (1988: 78)

16 See also the criticism developed by Anthony (1960), reprinted in this volume.

17 In order that SWM should lead to the enhancement of social welfare, monopolies and externalities should not exist.

18 On this point, see Anthony 1960, reprinted in this volume.

19 A large body of evidence suggests that investors diversify their portfolio holdings by much less than is recommended by normative models of portfolio choice. Investors are affected by the "home bias" since they prefer domestic equities (French and Poterba 1991) or stocks issued by companies located geographically close to them (Grinblatt and Keloharju 1999). Benartzi and Thaler (1998) find that when people do diversify, they do so in a naïve fashion.

20 Throughout this chapter the expression "provisions" is used to mean the funds provided to the firm-entity by various sources of financing, especially those related to the financial system such as debts and equities.

21 See the discussions at the beginning of this section and in previous sections.

## Bibliography

Allais, M. (1953) "Le comportement de l'homme rationnel devant le risque. Critiques des postulat et axiomes de l'école américaine," *Econometrica*, no. 21, 503–46.

Allen, F. (2001) "Do Financial Institutions Matter?," *Journal of Finance*, 56, 1165–75.

Anthony, R. (1960) "The Trouble with Profit Maximization," *Harvard Business Review*, 38, 126–34, reprinted in this volume.

Ardalan, K. (2000a) "The Academic Field of Finance and Paradigm Diversity," *Southern Business Review*, 26, no.1, 21–31.

——(2000b) "Development of the Academic Field of Finance: A Paradigmatic Approach," *Academy of Educational Leadership Journal*, 4, 44–79.

——(2004) "On the Theory and Practice of Finance," *International Journal of Social Economics*, 31, 684–705.

Arthur, W. B. (1990) "Positive Feedbacks in the Economy," *Scientific American*, 262, 92–99.

——(1995) "Complexity in Economic and Financial Markets," *Journal of Complexity*, I, 20–25.

——(1999) "The End of Certainty in Economics," in Clippenger, J. H. (ed.) *The Biology of Business*, San Franciso, Calif.: Jossey–Bass

Barney, J. B. (1986) "Strategic Factor Markets: Expectations, Luck, and Business Strategy," *Management Science*, 32, 1231–41.

Benartzi, S. and Thaler, R. (1998) "Naive Diversification Strategies in Defined Contribution Savings Plans," working paper, University of California–Los Angeles.

Berle, A. and Means, G. (1932) *Modern Corporation and Private Property*, New York: Macmillan.

Bettis, R. A. (1983) "Modern Financial Theory, Corporate Strategy, and Public Policy: Three Conundrums," *Academy of Management Review*, 8, 406–15.

Biondi, Y. (2005a) "The Firm as an Entity: Management, Organization, Accounting," Paper (Economia aziendale Dept.) no. 46, University of Brescia, August. Available at http://ssrn.com/abstract=774764

——(2005b) "A Generalisation of Investment Decision Criteria and its Application to the Firm and to Project-financing Partnerships," working paper, University of St Etienne, France, updated to October 2005.

Black, F. and Scholes, M. (1972) "The Valuation of Options Contracts and a Test of Market Efficiency," *Journal of Finance*, 27, 399–417.

——(1973) "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, 637–54.

Bline, D. M. and Cullinan, C. P. (1995) "Distributions to Stockholders: Legal Distinctions and Accounting Implications for Classroom Discussion," *Issues in Accounting Education*, 10, 305–16.

Brennan, M. J. (1995) "Corporate Finance Over the Past 25 Years," *Financial Management*, 24, 9–22

Canziani, A. (1987) "Sulle premesse metodologiche della rivoluzione zappiana," in AA.VV., *Scritti in memoria di Lino Azzini*, Milan: Giuffré.

Canziani, A. and Rondo Brovetto, P. (1992) "The Emergence of the Economics of the Firm in Continental Europe during the Twenties: Betriebswirtschaftslehre and Economia Aziendale as Methodological Revolutions," in Todd Lowry, S. (ed.) *Perspectives in the History of Economic Thought*, Aldershot, UK: Edward Elgar, 168–91.

Copeland, T. E. (1994) "Why Value Value?," *The McKinsey Quarterly*, no. 4, 97–109.

Copeland, T. E. and Weston, F. J. (1988) *Financial Theory and Corporate Policy*, Reading, Mass.: Addison Wesley.

Cutler, D. M., Poterba, J. M. and Summers, L. H. (1989) "What Moves Stock Prices?," *Journal of Portfolio Management*, 15, 4–12.

——(1990) "Speculative Dynamics and the Role of Feedback Traders," *American Economic Review*, 80, 63–68

De Long, J. B., Bradford, S. A., Summers, L. H. and Waldmann, R. J. (1991b) "Noise Trader Risk in Financial Markets," *Journal of Political Economy*, no. 4, 703–38.

De Long, J. B., Shleifer, A., Summers, L. H. and Waldmann, R. J. (1991a) "The Survival of Noise Traders in Financial Markets," *Journal of Business*, 64, 1–19.

——(1991) "Positive Feedback Investment Strategies and Destabilizing Rational Speculation," *Journal of Finance*, 45, 379–95.

Denrell, J., Fang, C. and Winter, S. G. (2003) "The Economics of Strategic Opportunity," *Strategic Management Journal*, 24, 977–90.

Dufrene, U. and Wong, A. (1996) "Stakeholders versus Stockholders and Financial Ethics: Ethics to Whom?," *Managerial Finance*, 22, 1–10.

Dybvig, P. and Ross, S. A. (1987) "Arbitrage," in Eatwell, J., Milgate, M. and Neuman, P. (eds.) *The New Palgrave: a Dictionary of Economics*, London: Macmillan, 100–106.

Ellsberg, D. (1961) "Risk, Ambiguity, and the Savage Axioms," *Quarterly Journal of Economics*, no. 75, 643–69.

Fama, E. (1980) "Agency Problems and the Theory of the Firm," *Journal of Political Economy*, 88, 288–307.

Fama, E. and Miller, M. (1972) *The Theory of Finance*, New York: Holt, Rinehart and Winston.

Findlay, M. C. and Williams, E. E. (1980) "A Positivist Evaluation of the New Finance," *Financial Management*, summer, 7–17.

Frankfurter, G. M.and McGoun, E. G. (1999) "Ideology and the Theory of Financial Economics," *Journal of Economic Behavior and Organization*, 39, 159–77.

French, Kenneth, and Poterba, James (1991) "Investor Diversification and International Equity Markets," *American Economic Review*, 81, 222–26.

Friedman, M. (1953) *Essays in Positive Economics*, Chicago, Ill.: University of Chicago Press.

Friend, I. (1973) "Mythodology in Finance," *Journal of Finance*, 28, 257–73.

Grinblatt, M., and Keloharju, M. (1999) "How Distance, Language, and Culture Influence Stockholdings and Trades," *Journal of Finance*, 56, 1053–73.

Jensen, M. C. (2001) "Value Maximisation, Stakeholder Theory, and the Corporate Objective Function," *European Financial Management*, 7, 297–317.

Jensen, M. C. and Meckling, W. H. (1976) "Theory of the Firm, Managerial Behavior, Agency Costs and Ownership Structure,' *Journal of Financial Economics*, 3, 305–60.

Jensen, M. C. and Smith, C. W. (1984) "The Theory of Corporate Finance: A Historical Overview," *Journal of Financial Economics*, 3, no. 4, 305–60.

Kahneman, D., Slovic, P. and Tversky A. (1982) *Judgment under Uncertainty: Heuristics and Biases*, New York: Cambridge University Press.

Kahneman, D. and Tversky, A. (1979) "Prospect Theory: An Analysis of Decision under Risk," *Econometrica*, 47, 263–91.

Lintner, J. (1965) "The Valuation of Risk Assets and the Selection of Risky Investments in Stock Portfolios and Capital Budgets," *Review of Economics and Statistics*, no. 47, 13–37.

Machlup, F. (1967) "Theories of the Firm: Marginalist, Behavioral, Managerial," *American Economic Review*, 57, March, 1–33.

Markowitz, H. (1952) "Portfolio Selection," *Journal of Finance*, 7, 77–91.

Merton, R. C. (1973) "Theory of Rational Option Pricing," *Bell Journal of Economics and Management Science*, 4, 141–83.

Miller, M. (1977) "Debt and Taxes," *Journal of Finance*, 32, 261–76.

Miller, M. H. (1986) "Behavioral Rationality in Finance: The Case of Dividends," *Journal of Business*, 59, S451–S468.

Miller, M. and Modigliani, F. (1958) "The Cost of Capital, Corporate Finance and the Theory of Investment," *American Economic Review*, 48, 261–97.

——(1963) "Corporate Income, Taxes and the Cost of Capital," *American Economic Review*, 53, 433–43.

Mossin, J. (1966) "Equilibrium in a Capital Asset Market," *Econometrica*, October, 768–83.

Mouck, T. (2000) "Beyond Panglossian Theory: Strategic Capital Investments in a Complex Adaptive World," *Accounting, Organizations and Society*, 25, 261–83.

Myers, S. C. (1977) "Determinants of Corporate Borrowing," *Journal of Financial Economics*, 5, 147–76.

Nicklisch, H. (1932) *Die Betriebswirtschaftslehre*, Stuttgart, Germany: Auflage.

Ross, S. A. (1978) "A Simple Approach to the Valuation of Risky Streams," *Journal of Business*, 51, 453–75.

——(2002) "Neoclassical Finance, Alternative Finance and the Closed End Fund Puzzle," *European Financial Management*, 8, 129–37.

——(2004) *Neoclassical Finance*, Princeton, N.J.: Princeton University Press.

Sauvain, H. (1967) "The State of Finance Field: Comment," *Journal of Finance*, 22, 541–42.

Sharpe, W. F. (1964) "Capital Assets Prices: A Theory of Market Equilibrium under Conditions of Risk," *Journal of Finance*, 29, 425–42.

Shiller R. J. (2000) *Irrational Exuberance*, Princeton, N.J.: Princeton University Press.

Shleifer, A. and Summers, L. H. (1990) "The Noise Trader's Approach to Finance," *Journal of Economic Perspectives*, 4, 19–33.

Shubik, M. (1993) "Accounting and Its Relationship to General Equilibrium Theory," in *Economic Notes*, by Monte dei Paschi di Siena, 22, no.2, 226–34, reprinted in this volume.

Simon, H. A. (1982) *Models of Bounded Rationality*, Cambridge, Mass.: MIT Press.

——(1983) *Reason in Human Affairs*, Stanford, Calif.: Stanford University Press.

Simon, H. A (1991) "Organizations and Markets," *Journal of Economic Perspectives*, 5, 25–44, reprinted in this volume.

Solomon, E. (1963) *The Theory of Financial Mangement*, New York: Columbia University Press.

Stauss, J. H. (1944) "The Entrepreneur: The Firm," *Journal of Political Economy*, 52, 112–27, reprinted in this volume.

Stiglitz, J. (1988) "Why Financial Structure Matters," *Journal of Economic Perspectives*, 2, no. 4, 121–26.

Thaler, R. H. (2000) "From Homo Economicus to Homo Sapiens," *Joumal of Economic Perspectives*, 14, 133–41.

Thaler, R. H. (ed.) (1993) *Advances in Behavioral Finance*, New York: Russell Sage Foundation.

Veblen, T. (1898) "Why Is Economics Not an Evolutionary Science?," *Quarterly Journal of Economics*, XII, July; also in T. Veblen, *The Place of Science in Modern Civilisation and Other Essays*, New York: Huebsch, 1919; reprinted New York: Russell & Russell, 1961, 56–81.

von Neumann, J. and Morgenstern, O. (1944) *Theory of Games and Economic Behavior*, Princeton, N.J.: Princeton University Press.

Weston, F. J. (1966) *The Scope and Methodology of Finance*, Englewood Cliffs, N.J.: Prentice-Hall.

——(1967) "The State of the Finance Field," *Journal of Finance*, 22, no. 4, 539–540.

Zambon, S. and Zan, L. (2000) "Accounting Relativism: The Unstable Relationship between Income Measurement and Theories of the Firm," *Accounting, Organizations and Society*, 25, 799–822.

Zappa, G. (1937) *IL reddito di impresa*, Milan: Giuffré.

——(1954) "La nozione di azienda nell'economia moderna," *Il Risparmio*, no. 8, 1255–78.

Zeckhauser, R. (1986) "Comments: Behavioral versus Rational Economics: What You See Is What You Conquer," *Journal of Business*, 59, S435–S449.

Zingales, L. (2000) "In Search of New Foundations," *Journal of Finance*, 55, 1623–53.

# 18 The corporate governance of the firm as an entity

## Old issues for the new debate

*Marc T. Moore and Antoine Rebérioux*

## Introduction

Corporate governance debates are primarily concerned with the allocation of power within listed companies, from a positive and a normative point of view. On both aspects, these debates have been structured, throughout the twentieth century, between managerialist and agency theories. The theory of "managerialism," as set out in the seminal work of Berle and Means (1932), stresses the inherent divergence of interest that occurs within widely held firms between corporate "owners" (shareholders) and "controllers" (managers), and the resultant "politicization" of the corporation as an object of public concern. In response to the perceived need to minimize this separation, "agency theory" suggests a range of market and contractual mechanisms aimed at bringing the incentives of managers into line with those of shareholders.

Over the last two decades, pro-shareholder mechanisms have become increasingly prevalent within corporate governance, in the United States and in the European Union. At first sight, this development would appear to signal the failure of managerialist theory, and the final victory of the agency perspective and its underpinning contractarian approach. In this chapter, we cast doubt on the above account. If traditional managerialist theory is somewhat misleading as a depiction of present day capitalism, this does not necessarily imply that shareholder primacy is "right." In fact, current evolutions show that the growing implementation of shareholder sovereignty has dramatic consequences, which are hardly explicable within the confines of the agency paradigm. The story becomes much more intelligible once we adopt a conception of the firm as an entity, rather than as a nexus of (complete or incomplete) contracts. The idea that corporate governance discussions should be based on a theory of the firm is obvious: it is hard to figure out how a firm is governed and how it should be governed without any conception of what is, precisely, a "firm." Therefore, this chapter intends to show that considering the firm as an entity allows us: (i) to understand the failure of shareholder primacy as a way to foster managerial accountability; and (ii) to reject shareholder primacy in favour of a more inclusive model of corporate governance. From this point of view, the normative conclusion put forth by Berle and Means (1932) still represents a powerful

insight: managerial accountability is best achieved through the promotion of common interest rather than through shareholder primacy.

The chapter is organised as follows. The first part offers an original account of corporate governance in the 1990s and the 2000s that runs against the predictive dimension of agency theory. In the second part, we highlight the limitations of this theory, and its underpinning contractarian logic, as a normative basis for corporate governance. Based on these findings, we set out in the third part a preliminary scheme for the implementation of Berle and Means' inclusive model of the firm within a market-based environment, taking insights from both the U.S. and the evolving European model of corporate governance.

## Positive assessment: whose interest does the corporation serve?

The multiplicity of factors that determine the allocation of power within listed companies gives rise to disputes as to which agents are the main beneficiaries of the corporate economy. We first present the debate between managerialist and agency theories before questioning the ability of both accounts to describe accurately the typical configuration of the 1990s and 2000s.

### *Managerialism versus agency theory*

In 1932, Adolf Berle and Gardner Means published what was to become one of the most influential and inspirational social-scientific works of the twentieth century. Berle was a corporate lawyer by trade, and Means an economist. The book in question is titled *The Modern Corporation and Private Property*. One U.S. historian has gone so far as to describe *The Modern Corporation* as "an ideological rational for New Deal planning, consumer activism, labor organizing, and financial regulation of the large corporation, indeed all of American capitalism" (Lichtenstein 2002). Berle and Means' thesis was concerned with the then growing economic and political phenomenon known as the widely held or "quasi-public" company. Unlike smaller, closely held or "private" companies, these larger companies were capitalised by the investment of finances from the private wealth of members of the public at large. The extraordinary nature and potential of the quasi-public company resided in the fact that, unlike smaller business units, in which a dominant shareholder or group of shareholders either managed the business personally or at least undertook some degree of active control over an appointed management team, the quasi-public company, at least in theory, exhibited a complete separation of ownership and control.

This was on account of the fact that the controlling managers of these quasi-public companies, unlike their counterparts in closely held firms, in many cases held a small or even negligible ownership stake in the firm, and therefore derived the main component of their income not from returns on company shares, but rather from a fixed salary in essentially the same vein as any other officer or employee of the company. The ownership of these firms, meanwhile,

was increasingly becoming vested in a multitude of small-scale individual investors, lacking both the resources and also the incentive to undertake effective control over the use to which management put their economic investment in the firm. This meant that, not only were the respective functions of share ownership, on the one hand, and business management, on the other, generally carried out by different and distinct persons, but the shareholders of the typical quasi-public company were generally neither willing nor able to exert effective control over management, so that, in effect, the function of overseeing management on a day-to-day basis became, somewhat perversely, the responsibility *of management itself.*

In *The Modern Corporation and Private Property*, Berle and Means (1932: 7) observed that "[t]he separation of ownership from control produces a condition where the interests of owner and of ultimate manager may, and often do, diverge, and where many of the checks which formerly operated to limit the use of power disappeared." Berle and Means distinguished the position of the modern shareholder from that of the pre-corporate owner-manager (see also Berle 1965, reprinted in this volume). The latter, although often taking no part at all in the day-to-day management of his business, was at least in a position to step in as and when he so chose, whether personally or through his subordinate managers on the "shop floor." The shareholder in the typical quasi-public company was, in contrast, "entirely quiescent," his position having "been reduced to that of having a set of legal and factual interests in the enterprise," without any corresponding effective powers over it, which, as a matter of fact, had become vested entirely in management (Berle and Means 1932: 112–13). Berle and Means therefore described corporate managers as "economic autocrats," whose ability to effectively perpetuate their own existence had promoted them to the position of "the new princes," assuming unchecked control over their "economic empires" (Berle and Means 1932: 116).

Furthermore, as Galbraith (1973) later noted, even to the limited extent that any shareholder or financial institution was sufficiently disposed to intervene from time to time in the operational affairs of the companies in which they were interested, any action that they took or demands that they made in this regard were inherently irrational, given the inability of these "outsiders" to acquire sufficient information or expertise to be able to properly pass judgement on the merits of managers' strategic decisions. Not only were shareholders physically detached from the day-to-day affairs of the business, but they were also excluded from what Galbraith termed the corporate "technostructure," which he defined as the collective body of corporate officers, including managers themselves who, by virtue of the supremacy that they, as a group, enjoyed over the base of scientific and technical skills, knowledge and expertise upon which the company's operations were dependent, enjoyed the exclusive capacity to command strategic control over all business affairs. In other words, Galbraith believed that, in the modern corporate economy, where operations were increasingly technical and specialised in nature, the "real" power within the large company rested with those that possessed the relevant *knowledge*, rather than the wealth, that

comprised the business, thereby excluding shareholders from the realm of effective corporate control.

In spite of its undoubted popularity in the first half of the twentieth century, Berle and Means' central descriptive claim, that the modern corporation no longer served the interests of its shareholders, was quick to attract strong opposition. Throughout the latter half of the century, a convincing academic counter-representation of the widely held company was developed under the rubric of "agency theory." This more "orthodox" branch of corporate scholarship, as developed in the works of the Chicago School economists Eugene Fama, Michael Jensen and William Meckling, adopted as its conceptual basis the contractarian paradigm, considering the firm as a self-determinative nexus of contracts linking together various individual input-providers (Fama 1980; Jensen and Meckling 1976; Easterbrook and Fischel 1991).[1]

In the context of corporate governance, the most significant achievement of agency theory was its capacity to reinvent, using a contractarian approach, the concept of the competitive market as a means of disciplining inefficient corporate managements. However, according to agency theorists, the most powerful discipline over management stemmed not from the product market, but, rather, from the market for the financial stock of companies themselves, which was forwarded as a much more compelling and therefore *relevant* institutional factor motivating continual improvements in managerial performance (Alchian 1969). Indeed, a liquid stock market is not only valuable as a medium through which firms must compete with one another to raise equity capital at low cost, but, more significantly, it is also a necessary prelude to the effective functioning of the market for corporate control and the associated disciplinary device of the hostile takeover (Manne 1965). Meanwhile, the detrimental effect of a low share price on managerial reputation provides corporate controllers with an ongoing incentive to prioritise the interest of shareholders, even in those instances where firm underperformance is not so severe as to warrant the initiation by an outsider of expensive hostile takeover proceedings (Fama 1980).

In so far, though, as the doctrine of shareholder primacy relies for its effective realisation upon the functioning of a liquid and efficient stock market, then it is necessarily supportive of legal and other institutional mechanisms that enable the continual publication of credible information on firm performance for the benefit of discerning investors. Indeed, without a reliable informational substructure, a liquid market in corporate securities is impossible. Within a complete market-based corporate governance system, reliable information on the firm is obtained through the interaction of two key groups: first, by the board of directors, which is formally appointed by and therefore controlled by shareholders; and, second, by external "gatekeepers," most notably financial auditors, securities analysts and ratings agencies. The role of the former body is to act as an independent supervisory panel situated between the shareholders' General Meeting and the management team, thus providing an "internal" point of surveillance over managers in the absence of direct shareholder monitoring. The latter actors, meanwhile, are vested with the responsibility of verifying the

honesty and relevance of financial information disclosed by the company's accounting reports, thereby reducing informational asymmetries between investors and insiders (agents in the company) so as to ensure the proper working of financial markets (Aglietta and Rebérioux 2005; Biondi 2004).

On the basis of this information, investors (shareholders) buy and sell securities, thus generating stock price movements, which in turn trigger either or all of the above market-disciplinary mechanisms. At the same time, managers' interests can be theoretically aligned with those of shareholders on an *ex ante* basis through the use of incentive-remuneration devices such as executive stock option plans. In accordance with the above mechanisms, then, shareholders enjoy the capacity not only to compensate for the separation of ownership and control at the level of the individual firm, but also to impose a more efficient form of control over management than could be achieved *via* direct supervisory oversight of individual companies. This is in view of the allegedly superior efficiency of the price signals of competitive investment markets in reflecting professionally acquired information on relative firm performance, coupled with the low transaction costs that shareholders face in disposing of their holding in shares of under-performing firms on a liquid market (Easterbrook and Fischel 1991).

We began this section by asking a question: whose interest does the corporation serve (at least in the Anglo-Saxon world)? We then reviewed two opposing answers to this question. According to the managerialist theory of Berle and Means, the separation of ownership and control within the modern corporation renders it subservient to the interests of managers, to the detriment of shareholders. According to agency theory, on the other hand, the corporation is rendered exclusively accountable to shareholders by virtue of market-based incentive and disciplinary mechanisms. The two answers are clearly different, yet they share one crucial assumption: they are both premised upon the existence of a *conflict* of interest between managers and shareholders, and this conflict represents the matrix, under each answer, of the so-called corporate governance "problem." The next section casts doubt on this shared point.

### *The paradox of the 1990s: a convergence of interest*

The success of the contractarian paradigm (of the agency perspective) may be appreciated in the academic sphere as well as in the political one, in so far as corporate governance is concerned: for the last twenty years, shareholder primacy has deeply influenced the evolution of corporate governance regulations and practices, in the United States and, to a lesser extent, in the European Union. The rights of (minority) shareholders got stronger everywhere, primarily through federal law in the U.S. and trans-national law in the E.U. (Cioffi and Cohen 2000). Besides, institutional investor activism has promoted best practices apparently akin to shareholder primacy. These investors seem to have succeeded, despite the diversification of their portfolios, in significantly increasing the sensitivity of the corporation to the shareholder primacy.

The growing success of a shareholder value oriented approach to managing a business can be observed on at least at three levels. First, the presence of independent non-executive directors, mostly in *ad hoc* committees (audit, nomination and remuneration), is now the rule rather than the exception: according to Finkelstein and Mooney (2003), outside directors accounted for 75% of directors in 2003 on the average board of firms in the Standard and Poor's (S&P) 500. Second, stock options are increasingly used as a remuneration device: whereas stock options accounted for less than 25% of the average S&P500 Chief Executive Officer (CEO) pay package at the beginning of the 1990s, this part has stabilised at around 50% since 1999 (Jensen and Murphy 2004). Last but not least, "Value-Based Management" – the use of management tools for establishing the creation of "shareholder value" – is now a common practice for listed companies (Cooper *et al.* 2000). Through these tools, constraining criteria of financial returns are imposed on firms (see p. 355). The competition among investment funds to attract collective savings is then transferred onto the companies, which are judged by these funds on the basis of their ability to meet the financial demands imposed on them.

By and large, the compliance of executive officers with shareholder primacy seems to be greater than it has ever been in the previous century. According to both the managerialist perspective *and* the agency view, this situation should be beneficial, in the first place, to stockholders. In effect, the wealth accruing to equity holders, through dividends, stock repurchase and increased market value has increased significantly since the 1980s. Conversely, managers should be the main loser: their ability to capture part of the profit stream should be reduced as well as their discretionary latitude in making business decisions. However, the case for this last assertion is not so good, leading to a striking paradox. Indeed, recent decades witnessed a huge rise in executive compensations, at least in the U.S. According to Holmström and Kaplan (2003) overall CEO compensation increased by a factor of six during the 1980s and the 1990s. Most of this increase took the form of incentive pay – primarily stock options. This process has resulted in a deepening of intra-firm inequalities, of which the *Business Week* pay executive survey gives an idea: in 1980, the average income of CEOs of the largest firms in the US was 40 times the average salary of a worker. In 1990, it was 85 times greater, and in 2003, it jumped to 400 times greater. A similar evolution is observable in the U.K.: the executive pay consultancy Independent Remuneration Solutions (IRS) has found that, since 1998, average CEO salaries have risen by 58% whilst total executive remuneration increased by 208%. In the same period, average earnings went up by 33% and retail prices by 15%, whilst the FTSE 100 index fell 13%.

These evolutions strongly suggest that the last decades have been characterised by a process of *convergence of interest* between stockholders and corporate officers. From a theoretical standpoint, it is crucial to note that this process does not fit with the basic assumption underlying the managerialist and the agency perspectives (see p. 351).[2] According to the latter, if shareholders succeed in aligning officers' incentives with their own personal financial interest, then an

optimal (first best or second best) contract is achieved. However, the current situation may hardly be considered "optimal," for at least two reasons. On the one hand, such an increase in officers' compensation raises serious concern *from a strict economic standpoint*: it is hard to explain on the basis of incentive factors alone, despite the effort made by some authors (see in particular Jensen and Murphy 2004). Rather, the most plausible account of this evolution is the occurrence of a process of rent extraction by corporate managers (Bratton 2005). On the other hand, confidence in financial markets has been seriously undermined by the major wave of high-profile corporate scandals and accounting irregularities that followed Enron and WorldCom's bankruptcies.

Such evidence indicates that a decline in managerial or corporate accountability took place during the 1990s, and may continue. We have already noted, however, that this decline seems by and large to profit shareholders. If this diagnosis is right, then the conclusion is clear-cut: other stakeholders do not benefit from current practices and evolutions in corporate governance, nor the firm as a (productive) entity. In the rest of this section, we offer an explanation for our main diagnosis – the fact that the growing implementation of shareholder primacy leads to a decline in corporate accountability. This decline, we argue, is the direct consequence of the intrinsic limitations of a mode of control relying solely on market-based solutions – that is, of a mode of governance where the stock market is the only valuation machinery for firms. To better understand this statement, an in-depth analysis of the functioning and implications of the most popular "Value Based Management" tool – the Economic Value Added (EVA) – may prove to be particularly useful: indeed, this metric, in both its informational and operational dimensions, reveals the logic of (stock) market control over listed companies.

"Value-Based Management" tools – and in particular the EVA metric – are said to offer to investors on the stock market the technical capacity to accurately assess business conduct. Indeed, the EVA metric is supposed to condense the complex sphere of information and contingencies determining the success of any one firm in its product markets into one or more general, all-encapsulating measure(s) of firm performance. The first function of EVA is therefore *informational*: it is considered to be the most relevant criterion for the prediction of stock market prices. The key point is not that value-based performance measures are comprehensive, in the sense that they reflect all or even nearly all of the information that shareholders would otherwise require to make a "rational" economic decision. Rather, and crucially, the point is that value-based performance measures are informationally *selective* mechanisms for evaluation of management, in that they possess the technical capacity to abstract, out of the complex field of economic and political contingencies determining the "right" strategic direction of the firm, the specific, definite and arbitrary performance yardsticks that are "relevant" to the evaluation of management by shareholders over any given time period. As a result, value-based performance measures promise to vest shareholders, particularly professional portfolio investors, with the capacity to compensate for the informational deficit that they encounter *vis-à-vis* management

consequent upon the separation of ownership and control. Interestingly, however, shareholder value achieves this end not by increasing the informational base that shareholders actually enjoy in assessing managers but, rather, by *reducing* the informational base that shareholders *need* enjoy in order to be able to make an allegedly "rational" judgement of management as such. Specifically, this minimal necessary informational base under current convention comprises the difference between financial profitability (the accounting Return On Equity) and the firm's cost of capital. Here we meet the second function of EVA, the *operational* one: EVA is set down as *the* management criterion for executives, who must seek to maximise the difference between the ROE and the cost of capital. This latter figure is no longer considered to be a consequence of the firm's productive and commercial operations, determined *ex post*. Rather, cost of capital is now a benchmark in itself, determined *ex ante*. The use of benchmarking thus provides financial investors with the ability to undertake a continuous and generalised comparison between listed companies.

Let us denote $R$ to be the net result, $D$ the book value of debts, $r$ their average costs, $EC$ the book value of equity capital, $k$ the equilibrium return on equity capital (or the cost of capital) as determined by the Capital Asset Pricing Model, $WACC$ the weighted average capital cost ($WACC = k.EC/K + r.D/K$) and $K$ the total book value of the assets ($D + EC$). The simplest expression of a company's EVA is then the following:

$$EVA = R - k.EC \tag{1}$$

By denoting $ROE$ the return on equity ($R/EC$), $ROA$ the return on assets ($R/K$) expression (1) rewrites:

$$EVA = (ROE - k).EC \tag{2}$$
$$= (ROA - WACC).K \tag{3}$$

These alternative expressions indicate the different (financial) strategies used to produce and increase EVA:

- the repurchase of shares ($\Delta EC < 0$), that increases the return on equity ($ROE$) – see (2);
- the asset-light strategy ($\Delta K < 0$), that automatically raises the return on assets ($ROA$) – see (3);
- the increase of the debt-to-equity ratio, when the cost of debt $r$ is below the cost of equity capital $k$. This decreases the WACC and therefore raises EVA – see (3).

These methods have been used extensively by officers of Enron, WorldCom and Ahold – some of the most representative corporate scandals of the 2000s. Through value-oriented financial engineering techniques, managerial wealth increases irrespective of, or even at the expense of, corresponding improvements

in firms' *productive* efficiency. Clearly, none of the previous methods are sustainable in the medium-to-long term. These are short-term strategies aiming at fostering financial returns beyond the market equilibrium. As such, they are highly risky and encourage bold innovations flouting acceptable standards of caution.

At this stage of the argument, it is important to note that the devices available for monitoring corporate officers are not necessarily strengthened by a strict, scrupulous implementation of shareholder value. The case of the board of directors is striking. According to institutional investors as well as shareholder primacy proponents, the *raison d'être* of the board of directors is to enable control *from inside* of the managerial team on behalf of distant stockholders, rather than the strategic assistance of executives in their business choices. Following this line, independence – as a way to prevent collusion between the controllers (board members) and the controlled (managers) – came to be a cornerstone of corporate governance reforms. If it is hard to give a precise content to the concept of independence, institutional investors need clear signs, visible from a distance. Among these signs, the absence of relationships with management is favoured. But as Roberts *et al.* (2005) note, such an approach towards independence tends to limit the involvement and engagement of non-executive directors in corporate affairs. In turn, this means a rather weak knowledge of the firm and its productive and commercial dynamics. As Roberts *et al.* (2005: 19) conclude: " ... the advocacy by institutional investors, policy advisors and the business media of greater non-executive independence may be too crude or even counter-productive."[3] The assessment of the board of directors offered by the doctrine of shareholder primacy is therefore paradoxical in that it advocates an increasing *exteriority* for this *internal* mode of control.

In a (pure) shareholder-oriented mode of governance, the role of the gatekeepers (rating agencies, securities analysts and external auditors) is then obviously crucial: once insider mechanisms of control – such as worker representatives or informed (internal) directors – have been dismissed, the gatekeepers *de facto* become the central (unique) supervising device. If they did fail in these high profile scandals, cognitive reasons may be at least as important as incentive concerns:[4] by nature, the gatekeepers are limited in their ability to evaluate the origins of corporate profits, which are intimately linked to the functioning and dynamics of the firm as an entity – they constitute the peculiar income to the firm stressed by Biondi (this volume). External surveillance devices, however sophisticated they may be, have intrinsic limits that point to the cognitive dimension of control and coordination. We mentioned Galbraith's (1973) argument that the power of corporate insiders derives from, and is justified by, the informational advantage that they enjoy over liquid shareholders (see p. 350). Any attempt to empower the latter constituency therefore comes up against this cognitive issue – a fact that shareholder primacy proponents seem to have bypassed rather too easily.

Let us sum up our main argument. If capital markets are now able to command results (through the generalisation of benchmarking), they are limited in

their ability to appreciate the way in which these requirements are met. This contributes to making managerial power *less* accountable: financial irregularities multiply and executive remunerations explode. Shareholder primacy fails exactly where it strives to succeed: it reinforces the discretionary power of managers rather than limiting it. We have argued that this failure highlights the limitation of a pure external (market-based) control – a notion that lies at the core of the doctrine of shareholder primacy in the context of a liquid stock market. According to this view, the firm is regarded as a standard financial asset, with one important consequence: a level of financial profitability is required *ex ante* (the cost of capital). Meanwhile, firms within the same risk category (as defined by financial analysts) are judged on a single dimension: their ability to overcome this benchmark, regardless of the specific circumstances of the business concern in question. This conception, however, denies the productive and cognitive dimension of the firm as an entity (see also Marzo, this volume). Interestingly, we find a similar oblivion in the contractarian paradigm, focused on the exchange rather than production process (see Weinstein, this volume). From this point of view, the inclination of the contractarian approach to advocate shareholder primacy should come as no surprise.

## Normative assessment: whose interests should the corporation serve?

Our analysis of contemporary corporate governance practices, marked by a serious drift in corporate accountability disguised behind the façade of compliance with shareholders' interests, logically leads us back to the classical question: whose interests should the corporation serve? Whereas we intend to demonstrate that Berle and Means' positive account of divergent interests between stockholders and officers represents in some sense an inaccurate analysis of the contemporary situation, we would argue that their normative account offers some powerful insights.

### *Berle and Means: the "institutionalisation" of the modern corporation*

Berle and Means' (1932) Book IV – the final one – opens with the following passage:

> The shifting relationships of property and enterprise in American industry ... raise in sharp relief certain legal, economic, and social questions which must now be squarely faced. Of these the greatest is the question in whose interests should the great quasi-public corporations ... be operated.
> (Berle and Means 1932: 294)

They identified two alternative answers, corresponding to two different doctrines: on the one hand, the doctrine of managerial sovereignty; and, on the

other, that of shareholder sovereignty. The *managerial sovereignty* doctrine takes cognizance of the concentration of power in the hands of the managers, observing that it is the result of a strictly contractual process: the shareholders have accepted loss of control over the company in exchange for greater liquidity. In other words, they have traded control for liquidity (Berle and Means 1932: 251). Consequently, the shareholders can no longer legitimately demand control over the company, so that ultimate power of direction over the firm rests with managers. Berle and Means (1932) expressed concern about this approach on the basis that it gives almost dictatorial power to the managers, whom they described as "the new princes" (see p. 350). Although Berle and Means regarded the *shareholder sovereignty* doctrine to be a better (or, at the very least, *aless bad*) solution, they were not especially enthused by it either, precisely because it refuses to acknowledge the trade-off between control and liquidity. A careful reading of subsequent writings by Berle – and in particular of "The theory of enterprise entity" (Berle 1947, reprinted in this volume) – suggests another reason for these authors' unease with shareholder primacy as a guideline for corporate power. This reason is the following: shareholder primacy, at least as a legal doctrine, tends to "hypertrophy" the corporation (as a legal device with artificial personality) to the detriment of the firm (as an enterprise, i.e., productive, entity). It is precisely against this "bias" of the "corporate theory" that Berle (1947: 344, reprinted in this volume) has developed the "theory of enterprise entity":

> The divergence between corporate theory and the underlying economic facts has occasioned a variety of problems (dealt with *ad hoc* by the courts) . . . It is the thesis of this essay [t]hat the entity commonly known as "corporate entity" takes its being from the reality of the underlying enterprise, formed or in formation.

This article is of foremost importance, for it suggests that the dismissal of shareholder primacy is, at least in Berle's mind, rooted partly on consideration of the intrinsic economic nature of the firm (see also Manfrin, this volume). Put differently, to focus on the productive dimension of the firm rather than on its legal or financial aspects supports the idea that corporate governance should not deal solely with shareholders.

Let us now return to Berle and Means' position concerning the accountability of corporate managers. This position is briefly presented in their very last chapter. This chapter begins with a long quotation from Walther Rathenau, industrialist, statesman in the Weimar Republic and social theorist, describing the German conception of the public limited company in the following terms: "The depersonalization of ownership, the objectification of enterprise, the detachment of property from possessor, leads to a point where the enterprise becomes transformed into an institution which resembles the state in character" (Berle and Means 1932: 309). Likewise, in the new introduction for the 1967 edition of the *Modern Corporation and Private Property*, Berle wrote: "There is an increasingly recognition of the fact that collective operations, and those

predominantly conducted by large corporations, are like operations carried on by the state itself. Corporations are essentially political constructs" (Berle and Means 1932: xxvi).

Both quotations shed light on the distinction between two antagonistic logics. According to the logic of *ownership*, the (legal) world is divided between owners (legal persons, whether human or non-human) and objects of ownership. The owner of an object has "subjective" power over that object, which means that he has the right (the power) to do whatever he wants with it under the law (Robé 1999). Note that shareholder sovereignty and managerial sovereignty both analyse the corporation through this logic: the company is an object of ownership. The difference is the identity of the owners. According to the doctrine of shareholder sovereignty, the only legitimate owners are the shareholders. According to the managerial sovereignty thesis, the ownership has been traded off in favour of liquidity, so that managers are the real "owners." In contrast, the logic of *institution* dictates that the holder of power should not be free to exercise it in his interest (subjectively), but in the interests of those affected by it. The reference to the state in Rathenau's and Berle's quotations is significant on this level: the distinctive feature of a non-totalitarian state resides in the fact that the concentration of power within the state apparatus, necessary for its efficiency, is counterbalanced by limits placed on that power. The exercise of power is subjected, by means of various procedures, to the will of the people. Hence, the idea defended by Berle and Means is that the liquidity of stock markets calls for a rethinking of the nature of power within large companies. The firm is no longer an object of property, but an *institution* that must be governed as such. If the corporation is an institution – meaning that subjective interest should not be a guideline for the exercise of power – then it is necessary to set limits on managerial power to ensure that it is exercised on behalf of the company's constituents: shareholders, certainly, but also workers and, even further, the communities in which these companies thrive. Managers should not be accountable solely to the shareholders; they must be made accountable to all the stakeholders in the firm. *The Modern Corporation* therefore ends with a plea for management that would be a "purely neutral technocracy" (p. 312). Ultimately, whereas the agency perspective seeks to *minimize* the separation of ownership and control, Berle and Means offer to *exploit* it in order to enhance the role of public concern in capitalism.

### *The contractarian approach: towards a rejection of shareholder primacy*

The rejection of the concept of ownership, as applied to the business firm, is a standard assumption of the contractarian approach in law and economics (see, for example, Fama 1980). However, the similarity with Berle and Means is only superficial. This rejection is bound to a conception of the firm as a "nexus of contracts." By definition, one cannot possess a contract (or contracts) as one can possess a standard asset. However, the core of the "ownership conception"

remains: as argued earlier, the agency model confers upon shareholders subjective power over the corporation – even if this subjective power is *de facto* limited by the opportunism of corporate executives. The substance of the agency model is unambiguous: an efficient corporation is a corporation where shareholders are able, through a diversity of mechanisms, to impose their subjective interests. Shareholders are not depicted as owners, but as sovereigns. The implications, as far as corporate governance is concerned, are basically the same: managers and directors should be accountable solely to stockholders. Accordingly, exclusive control by stockholders over the board of directors is necessary.

It is therefore remarkable that the most recent works on corporate accountability inside the contractarian approach tend to give credence to the idea defended, over 70 years ago, by Berle and Means (1932). For that reason, the articles by Zingales (1998 and 2000) and Blair and Stout (1999) are of particular importance and deserve careful examination. From a methodological point of view, these articles put forward the notion of contractual incompleteness – so that Zingales (1998) writes of the "incomplete contracts approach to corporate governance."

The hypothesis of contractual incompleteness lies at the heart of the contemporary theory of the firm. It is one of the foundations of transaction cost theory, pioneered by Williamson (1975; 1985), and of modern property rights theory, developed by Grossman and Hart (1986). Both these approaches explore the way in which parties to a transaction secure their reciprocal investments when contracts are incomplete. In this context, protection of specific, non-redeployable investments cannot be achieved beforehand by the establishment of a contract providing for every possible contingency. Consequently, the parties to the contract are led to establish institutional devices, enabling them to appropriate a share of the organisational quasi-rent as a return on their investment. When applied to corporate governance, this schema considers rights on the board of directors as a tool for securing investments.

This path was first explored by Williamson in Chapter 12, entitled *Corporate Governance*, of his seminal 1985 book *The Economic Institutions of Capitalism*. His argument is taken up and furthered in two articles, one by Williamson and Bercovitz (1996), the other by Romano (1996). These works recognise that shareholders are not the only risk-takers within the firm. In particular, the increase in the specificity of human capital constitutes a risk-taking factor for the workforce: workers' payoff depends on the future distribution of the quasi-rent generated by the investment in human capital, which is fundamentally uncertain. This risk is all the stronger as the specificity of capital, in other words its non-redeployable nature, places employees in a disadvantageous position at the time of (re-)negotiation of the allocation of the quasi-rent. Reflection is thus focused on the measures capable of efficiently protecting those parties which incur the greatest risk (shareholders and employees), whereas contracts are incomplete. These authors reach the following conclusion: shareholders should be protected through rights of control over the board of directors. As for

employees' investments, they should be secured by means of various devices: a pre-defined system of promotion, severance packages and procedures for settling internal disputes. Employee participation on the board of directors, however, is not envisaged. One may be surprised by this asymmetry between the treatment of shareholders and that of employees: in one case, the recognition of a weakness gives the right to control; in the other, it gives the right to protection against the arbitrary nature of decisions. The conclusion of Romano (1996: 293) is clear: "Transaction cost economics offers no analytical support for expanding board representation to non-shareholder groups, and indeed, cautions against such proposals." The German model, in which employee representatives sit on the supervisory board,[5] is deemed to be inefficient.

The work of Zingales (1998; 2000) and of Blair and Stout (1999) has developed the incomplete contract approach to corporate governance pioneered by Williamson. The role of human capital, through specific investment, in the generation of organisational quasi-rent is fully acknowledged. Besides, Zingales (1998; 2000) and Blair and Stout (1999) observed that the quasi-rent created by the firm derives from the pooling of complementary factors of production, in the form of tangible, and also intangible, human and financial capital. Compared with the work of Williamson, more emphasis is therefore placed on the synergies that come into play between the investments of the different stakeholders. The firm is thus conceptualised as a "nexus of specific investments." The allocation of rights of control over the entity thus created plays a decisive role, in that this allocation will determine how the value created is divided up within a framework of contractual incompleteness. Consequently, each stakeholder will be more or less motivated to commit to the firm, and this will influence the very level of the organizational quasi-rent.

Taking into account the complexity of the relationships formed between the different stakeholders, Zingales (1998) and Blair and Stout (1999) propose a solution that moves away from the doctrine of shareholder value: the stakeholders should delegate their powers to an independent third party – the board of directors – whose objective is to serve the best interest of the constituted entity. In this context, the directors are no longer simply the agents of the shareholders; their fiduciary duties must be exercised towards the whole firm. Thus, the productive capital of the firm must be managed in the interest of the firm itself. This point is new compared with the work of Williamson, for whom the role of the board of directors was to serve the interests of the shareholders. In short, the primacy of shareholders is partially challenged, in order to foster firm-specific investments.[6] Zingales (2000: 1645) goes one step further in an article of a very prospective nature entitled *In Search of New Foundations*: "In the current environment, where human capital is crucial and contracts are highly incomplete, the primary goal of a corporate governance system should be to protect the integrity of the firm, and new precepts need to be worked out."

Examination of the contributions of Zingales (1998; 2000) and of Blair and Stout (1999) therefore brings out a remarkable principle: the stronger the emphasis on contractual incompleteness, the more corporate accountability is

extended. Put differently, it appears that the rejection of shareholder primacy is linked to a (theoretical) recognition of the incomplete nature of contracts. Zingales (2000) thus calls for reflection on new principles of governance in light of the current situation in which contracts are "highly incomplete" (see p. 361). From a mode strictly centred on shareholders (Williamson 1985), principles of governance were first extended to the management of productive capital in a common interest (Zingales 1998; Blair and Stout 1999), before finally embracing protection of the "integrity of the firm" (Zingales 2000).

It should be noted, however, that this principle, by which the widening of the field of incompleteness is accompanied by an extension of corporate accountability, raises deep theoretical problems. The widening of the field of incompleteness progressively reduces the validity of the contractual analysis on which the work of Zingales and of Blair and Stout is founded. To say that contracts are incomplete is to acknowledge, by definition, that the "off-contract" plays a role in coordination. Broadly speaking, the "off-contract" is constitutive of all the social regularities or forms (routines, norms, conventions, legal rules etc.) that underlie the productive process and are not the direct outcome of voluntary agreement. Weinstein, in this volume, offers a survey of the approaches to the firm – such as the resource-based and evolutionary perspectives – that focus on those regularities, taking into account the cognitive dimension of intra-firm coordination. As the incompleteness of contracts increases, the scope of contractual analysis tends to shrink (Favereau 1997), whereas the analytical consistence of those cognitive perspectives increases. This difficulty is recognised implicitly by Zingales (1998: 502), for whom "at the current state of knowledge the [incomplete contracts approach to corporate governance] lacks theoretical foundations." The author adds by way of conclusion: "Without a better understanding of why contracts are incomplete, all the results are merely provisional."

The argument developed in this section may be summarised as follows. The contractarian approach strongly advocates shareholder primacy, mostly through the agency conception of the firm. None the less, there are some exceptions that severely qualify or reject this model of corporate governance. On closer scrutiny, these exceptions share a common assumption: the incomplete nature of contracts is recognised as a crucial feature of the business firm. In turn, it is essential to note that such recognition ultimately amounts to concluding that intra-firm economic coordination is partially (or essentially) outside the realm of contractual order. This is precisely the core meaning of the theory of the firm as an entity. *In fine*, acknowledgement of the *specificity* of productive activity and dynamics leads to a rejection of shareholder primacy, for at least two reasons: the need for internal control mechanisms (first section) and the need to foster firm-specific investments (second section).

## The governance of the firm as an entity

Having highlighted the receding relevance of the classical theory of the firm,[7] Berle and Means had paved the way for the development of a new

social-scientific model of the enterprise, which took account of the revolutionary changes brought on by the development of the quasi-public corporate form. As discussed above (see p. 359), the new conception of the company which Berle and Means (1932: 312) suggested was that of a "purely neutral technocracy" where management would be required to resolve its freedom from market-institutional pressures in order to further some set of commonly agreed politico-economic goals. It is at this point, though, that we come face to face with a vexing paradox in Berle and Means' work. Whilst their apparent goal in writing *The Modern Corporation and Private Property* was to challenge the financial shareholder's exclusive status as the legal beneficiary of corporate managerial decisions, the authors' very highlighting of the separation of ownership and control contrarily served to earmark the shareholder–managerial relation as *the central focus* for future Anglo-Saxon corporate governance scholarship. In this way, then, *The Modern Corporation* had the unintended consequence of providing a conceptual frame of reference for future academic debate on the issue of how to minimise (rather than actually exploit) the so-called separation of ownership and control within the modern corporation. In part, this was due to the early caution shown by Berle (1932) himself in stressing the need for courts and legislators to maintain the legal principle of shareholder primacy in the absence of a robust regulatory scheme for making company directors directly answerable to the community as a whole. However, cause must also be attributed to the failure of subsequent corporate scholars to use Berle and Means' work as the foundation for a thoroughgoing conceptual remodelling of the business enterprise. Had such a reformulation of the modern corporation been carried out, it may have been capable of providing normative content to Berle and Means' (1932: 313) fledgling proposal for a director's legal duty of "economic statesmanship."

To this end, our aim in this final section is to put forth suggestions on how Berle and Means' conception of the company as a publicly oriented institution might be effectively implemented within a contemporary market-based corporate governance system. We consider first, as a regulatory means towards this end, the anti-takeover legislation that has been developed in recent decades within the majority of U.S. state corporate law systems. In vesting managerial boards with formal discretion to block hostile takeover bids, such provisions effectively promote the integrity of the productive enterprise above any conflicting considerations of shareholder value. Moreover, anti-takeover legislation serves to earmark managerial discretion, rather than shareholder self-interest, as the basic institutional determinant of corporate prosperity. At the same time, though, it is essential to recognise the ultimate futility of relying upon managerial discretion alone to bring about a more inclusive system of corporate governance within an Anglo-Saxon context, given the underlying alignment of the managerial interest with that of stockholders (see p. 353). Accordingly, we argue, anti-takeover provisions must be supplemented by effective procedural mechanisms for ensuring the influence of key *non*-shareholder interests at the

heart of the corporate governance process, at least if the law is to have any meaningful effect in breaking down the organisational dominance of influential stockholders and managers. In this latter regard, we therefore rely on the evolving European model of corporate governance and, in particular, its central theme of promoting the participation of employees in corporate decision-making *via* formal rights to information, consultation and (to a limited extent) representation on the board of directors itself.

### *The interest of the company: insights from U.S. anti-takeover legislation*

In the United States at present, the clearest doctrinal embodiment of the entity conception of the firm is the collection of controversial non-shareholder constituency laws that regulate the conduct of takeover bids at state level. In the 1980s, following widespread political and popular unrest with the social consequences of finance-driven takeover activity, many American states implemented radical reforms in their corporate statutes in an attempt to counteract this potentially harmful phenomenon. The common (and intended) effect of these amendments was to vest managements of target companies with the discretion to block hostile takeover attempts, using defensive tactics if necessary, where the directors felt in their good-faith judgement that the takeover would overall have a negative effect upon the corporation and its constituents (Parkinson 1993). To date, 43 out of America's 51 state corporate law systems[8] contain an anti-takeover provision of this sort,[9] which typically exhibit a number of salient features. The defining characteristic of such rules is the explicit discretion which they afford a target company's board of directors to consider the consequences of the takeover for "any or all groups affected by such action, including shareholders, employees, suppliers, customers and creditors," together with "communities in which offices or other establishments of the corporation are located" (Pennsylvania Consolidated Statutes, Title 15, § 511(d)(1)). Some versions of the rule even go so far as to specify "[t]he economy of the state and nation" (Ohio Revised Code, § 1701.59(E)(2)) as a legitimate subject of managerial concern.

A further typical feature of such provisions is the licence that they afford managements to take into account "[t]he long-term and short-term interests of the corporation and its stockholders, including the possibility that these interests may be best served by the continued independence of the corporation [from its potential acquirer(s)]" (General Laws of Massachusetts, Part I, Title XXII (Corporations), Chapter 156B, Section 65). One version of the rule adds, as a further ground for consideration, the general value of "the stable, long-term growth of domestic public corporations" (South Dakota Domestic Public Corporation Takeover Act, § 47-33-2(3)), while another pays specific regard to "benefits that may accrue to the corporation from its long-term plans" (Pennsylvania Consolidated Statutes, Title 15, § 511(d)(2)). Most extensively in this regard, the South Dakota Domestic Public Corporation

Takeover Act contains an express Declaration of Public Policy, which (*inter alia*) asserts that:

> '[t]akeovers of publicly held corporations are . . . frequently financed largely through debt to be repaid in the short term through changes in the operations of the target corporation, by the sale of substantial assets of the target corporation, and other means. In other states, such takeovers have impaired local employment conditions and disrupted local commercial activity. These takeovers . . . may undermine the state's interest in promoting stable relationships involving the corporations that it charters' (§ 47-33-2(5)).[10]

A final and complementary feature of non-shareholder constituency laws is their general discouragement of directors, when deciding upon the merits of a takeover proposal, from regarding the interest of any one or more particular constituent groups (e.g. shareholders) as "a dominant or controlling factor" (Indiana Code 23-1-35-1(f)) in their decision *vis-à-vis* the interest of the productive enterprise as a whole.

These provisions represent an *ex ante* statutory guard against the potentially degenerative effects of finance-led acquisition policies. Moreover, by recognising the fact that hostile takeovers very often place the dual interests of the corporation and its shareholders starkly into conflict with one another, U.S. anti-takeover legislation represents an authoritative refutation (at least within the takeover context) of the orthodox contractarian assumption that what is good for shareholders is, by implication, also good for enterprise and society (Coates 1989; Millon 1990).[11] However, while the end result of these provisions would appear to be that of reducing the organisational dominance of shareholders within the U.S. corporate governance arena, the U.S. anti-takeover machinery is, in reality, emasculated in its effect by the practical reality of how public corporations are actually run within America. Indeed, despite their apparent theme of politico-economic "inclusivity," American anti-takeover protections are ill-equipped to ensure that managements are made any more answerable to vulnerable non-shareholder groups than would be the case under an "orthodox" shareholder-oriented understanding of the company's interests. This is due to the fact that these provisions are underlain by an assumption of management autonomy, which can be exploited in order to ensure that shareholders' interests do not ride rough-shod over those of other participants whenever the shareholder and general corporate interest come into conflict with one another.

As logical as the foregoing course of reasoning may be, however, it is premised upon one erroneous factual premise: management is not "autonomous" in either of the above respects, but is in fact heavily coloured in its day-to-day decision-making by the same (or at least very similar) financial motivations as the shareholder. This point is not in need of any further explanation, but rather is borne out by the arguments put forth earlier (see pp. 352–57). It therefore follows that a board of directors, if left to its own devices, will not be necessarily prone to regard a financially driven takeover attempt as in conflict with their

own material interest. Indeed, not only will the resultant rise in firm share price probably benefit the incumbent managers themselves via incentive-remuneration schemes, but senior management will also likely enjoy a significant control premium, whether in the form of a lucrative (albeit less senior) office within the reorganised company (Manne 1962) or, at the very least, a substantial "golden parachute" payment upon termination of office (Coffee 1988).

Accordingly, while the U.S. directors' fiduciary duty (as typically formulated within state corporate law systems today) is sufficiently designed so as to afford legal protection to management (and, indirectly, other corporate participants) in the face of acquisition attempts that are potentially harmful to the incumbent board personally, it is not designed to afford direct protection to vulnerable non-shareholder groups in the more likely event that such a project will reap reciprocal gains for shareholders and senior executives alike. As a result, American corporate law cannot stop management from sanctioning a potentially harmful acquisition attempt that promises side-benefits for them personally, nor indeed any other form of financial restructuring (e.g. a downsizing project, share buy-back plan, or debt-for-equity restructuring) which serves to favour the financial wellbeing of the shareholder–managerial coalition at the expense of the integrity of the productive enterprise as a whole.

### *Empowering non-shareholder constituents: insights from the European model*

This is not to say, however, that existing U.S. anti-takeover legislation is incapable of providing at least a basic guide as to how Berle and Means' "neutral technocracy" conception of the company might be effectively implemented today within a market-based corporate governance system. Indeed, not only do many of the statutory provisions referred to above emphasise the socio-economic value of the company's stable growth and long-term strategic plans, but many also recognise the likelihood of there being conflict between, on the one hand, the financial interest of shareholders, and, on the other, the interest of the enterprise entity in the wider sense. In this way, American formulations of the corporate interest are at the very least capable of providing the seeds for a new publicly orientated understanding of the doctrine (Teubner 1994), which might represent a countervailing influence to the dominant mantra of the EVA metric and other managerially-enriching performance measures. However, given the position of the shareholder–managerial coalition as a significant vested interest in the corporate wealth-distributional process, coupled with the proven discretion enjoyed by management in allocating corporate returns, it becomes clear that fundamental changes in actual decision-making structures are a further necessary prelude to the legal implementation of the entity conception of the firm within an Anglo-Saxon context. In particular, it is submitted that the effective "institutionalisation" of shareholder–managerial prerogative in the Berle–Means sense is contingent not only upon the vesting of the board with ultimate rights of veto over hostile takeovers and other financial restructurings,

but also upon the installation of truly independent non-shareholder interests within companies' internal decision-making mechanisms (Stone 1975; Nader *et al.* 1976; Teubner 1985; Parkinson 1993).

In this latter respect, the developing "European model" of corporate governance stands out as a fledgling model of good practice for its Anglo-Saxon counterpart (Rebérioux 2002). Indeed, a dominant theme apparent in recent E.U. corporate and labour law legislation has been the Community-wide expansion of mandatory mechanisms for management–labour dialogue on important issues relating to the strategic and technological development of the employer undertaking. These measures include the E.C. Directive of 2002 establishing a general framework for informing and consulting employees in the European Community (Directive 2002/14/EC – "the ICE Directive"), which stresses (in its recital) the need for Member States to take efforts to ensure "that all citizens benefit from economic development" through the promotion of "social dialogue between management and labour." To this end, the ICE Directive demands generally that the management of a large undertaking, employing at least 50 employees in any one Member State, initiates procedures at the request of the workforce to enable the periodic provision of relevant information to elected employee representatives for purposes of consultation. Specifically, Article 4 of the Directive obliges management to disclose information on "the recent and probable development of the undertaking's activities and economic situation," together with the "situation, structure and probable development of employment within the undertaking," and "any anticipatory measures envisaged . . . where there is a threat to employment within the undertaking." In addition, management must report to employees on "decisions likely to lead to substantial changes in work organization or in contractual relations."

The requirements of the ICE Directive are backed up by the supplementary provisions of the European Works Council Directive (Council Directive 94/45/EC – "the EWC Directive"). Unlike the former scheme, the EWC Directive applies specifically to "community-scale undertakings," defined as those with 1,000 or more employees, at least 150 of whom are employed in two or more different E.U. Member States (Article 2). In essence, this latter Directive requires management (at the employees' request) to create the conditions for the setting up of a special European Works Council ("EWC"), comprised of three or more employee representatives, at least one of whom should represent each Member State in which the undertaking carries on its operations ("Subsidiary Requirements"). The specific purpose of the EWC scheme is to enable the provision of information by management to the workforce, and subsequent management–labour consultation, focused upon those transnational matters concerning the Community-scale undertaking as a whole, or at least two of its establishments situated in different Member States. As to the precise nature of the information to be provided and discussed, the Subsidiary Requirements of the EWC Directive list an extensive range of strategic and financial issues, including (*inter alia*): the undertaking's structure, economic and financial situation; the probable development of the business and of production and sales; the introduction of

new working methods or production processes; and transfers of production, mergers, cut-backs or closures of undertakings.

From the perspective of the argument at hand, the above provisions would at first sight appear to represent a valuable step in the right direction, in so far as they provide for the involvement of employees at a relatively early stage in the corporate decision-making process. This should, in theory, enable some degree of non-shareholder input into important strategic decisions (e.g. mergers or major financial restructuring projects) on an *ex ante* basis, thus vesting employee representatives with the formal role of "policing" controversial exercises of managerial prerogative. A notable contrast can be drawn here with the traditional structure of industrial relations in the United Kingdom, in which labour interacts on a primarily conflictual basis with management by "collective bargaining" via the channel of trade union representation, thereby perpetuating a view of employees as being situated "external" to the enterprise, with shareholders and managers enjoying exclusive "insider" status (Kahn-Freund 1956; Deakin and Morris 2005). Furthermore, both Directives contain an equivalent Article 9 provision, requiring consulting parties to "work in a spirit of co-operation and with due regard for their reciprocal rights and obligations, taking into account the interests of both the undertaking and the employees." The resultant expectation of reciprocal dialogue in discussions might better encourage a process of organisational "learning" on the part of management and labour representatives, thereby progressively integrating these parties' countervailing conceptions of the "good" of the company on a decision-by-decision basis (Teubner 1985).

Additionally, over the last couple of years, the subject of reforming the composition of company boards has been subject to some renewed interest on a pan-European level. The catalyst for this moderate resurgence of the "industrial democracy" debate was the introduction in 2001 of the long-awaited E.U. Directive on worker involvement in the European Company (Directive 2001/86/EC). In essence, this Directive makes provision for the involvement of employees in strategic decision-making, at board level, in any business that is registered as a Societas Europaea (SE) or European Public Limited-Liability Company. The SE is a specialist corporate form available to transnational undertakings conducting business in more than one E.U. Member State, aimed principally at reducing the transaction costs involved in effecting mergers and other reorganisations between companies subject to separate domestic legal regimes (see Regulation (EC) No 2157/2001 – "the European Company Statute"). In theory, use of the SE form provides such enterprises with the convenience of a uniform legal structure upon registration, applicable across the Community as a whole so as to avoid the inconsistency and confusion that results from different companies within the same undertaking being subject to differing sets of company law rules depending on their particular state of registration (Davies 2003; Deakin and Morris 2005).

Of relevance for present purposes are the provisions of the supplementary Directive on worker involvement noted above. In its original guise in the 1970s,

the European Commission's blueprint for worker involvement in the European Company proposed a mandatory system of worker representation at board level effective within SEs across the Community as a whole. The suggested board composition for SEs was based loosely on Germany's "two-tier" board model, featuring equal representation for shareholders and employees on an "upper" supervisory tier, which would be vested with the responsibility of appointing and overseeing a "lower" managerial board (Davies 2003). The employee representation requirements as set out in the final 2001 draft of the Directive are less stringent in nature, due in part to the opposition shown towards earlier drafts by some Member States, in particular Spain and the United Kingdom (Rebérioux, 2002). Nevertheless, Article 13 will have the effect of imposing mandatory employee representation requirements upon SEs registered in states with traditionally shareholder-oriented corporate law systems (e.g. the United Kingdom), in the event that that SE is formed as part of a joint venture involving a company whose "host" state already has such a system in force (e.g. Germany or the Netherlands). The potential effect of this provision, in encouraging the (limited) extension of "Rhineland" board structures into the Anglo-Saxon corporate governance systems of Britain and Ireland, both replies to and contradicts Hansmann and Kraakman's (2002) influential prediction that competitive and ideological forces are compelling international convergence towards the latter type of model.

Although the above developments are admittedly limited in scale, their symbolic value is undeniable. They constitute a new corporate design – purely European – in the global market, that is a coherent alternative to the Anglo-Saxon model. Indeed, the (limited) pan-European spread of industrial democracy within corporate governance exemplified by the above developments, together with their underpinning "social" rationale, breeds hope for a viable ideological alternative to the Anglo-Saxon orthodoxy of value-based management and its underpinning contractarian logic. In particular, the above developments could be regarded as indicative of a fundamental paradigm shift in the international corporate governance debate itself, inspired at root by two crucial contemporary developments explained above: first, the increasing convergence of the shareholder and managerial interests brought upon by value-based performance measures and related incentive-remuneration schemes (see p. 353); and, second (and correspondingly), the growing divergence between, on the one hand, the financial interest of shareholders (both managerial and non-managerial), and, on the other, the integrity of the productive enterprise entity together with the welfare of its key non-shareholder participants.

Against this backdrop, the (limited) institutionalisation of the views of the workforce within companies' decision-making procedures can be portrayed as a structural safeguard against the misalignment of financial and productive interests. To this end, labour representatives will be expected to challenge the purported integrity of controversial managerial statements by reference to key financial and non-financial information pertaining to both the company's financial performance and its recent industrial operations. The main responsibility

of labour representatives in this novel regard is to verify the honesty and justifiability of any managerial claim that a proposed corporate restructuring project, imposing costs upon vulnerable non-shareholder interests, is indeed motivated by "economic necessity," "the genuine needs of the business," or whichever variant of these terms is deployed by management to provide some moral vindication for the course of action in question. For example, the management of a company that decides to dismiss a large number of employees, or sever important links with long-standing suppliers, would find it difficult to justify the taking of that step if it is subsequently shown that, in the same or following financial years, that firm returned a considerable amount of its potential investment funds to shareholders via a large-scale stock repurchase. Likewise, a company that undertakes a significant downsizing project whilst, at the same time, paying exorbitant stock-based remuneration to its senior executives, would again face the need to justify any claim that it is motivated by genuine entrepreneurial considerations as opposed to mere shareholder–managerial control over the firm's wealth-distributional process. In such instances, the responsibility should fall on the company's management to reflect upon the reasons for its proposed policy, both by collecting (and disclosing to labour representatives) relevant financial and non-financial information, while also engaging in genuine reciprocal dialogue with employee representatives (whether at board or Works Council level) centred around the common aim of determining whether the proposed changes are indeed merited by competitive and/or technological factors.

## Conclusion

It may therefore be the case that, within the current cultural climate of "shareholder value," the corporate governance structure that is best capable of re-aligning the dual financial and productive dimensions of the enterprise entity will be derived from a combination of certain features of both the American and European models. The corporate governance of the firm as an entity, it is submitted, must be based at root upon the managerialist conception of the interest of the company characteristic of most U.S. state corporate law systems, which formally vests the board with discretion to prioritise the continuing integrity of the productive enterprise in the face of finance-led acquisition policies aimed centrally at the creation of shareholder value. However, in recognition of management's non-independence from the shareholder interest, our suggested model of corporate governance will also borrow heavily from the evolving European "social" model of corporate governance. In this article, we forward the European model as an institutional blueprint for the incorporation of employee representatives into the corporate governance process as a formal "check" on managerial decision-making discretion. In this way, labour has the potential to represent a locus of countervailing power *vis-à-vis* the dominant shareholder–managerial interest base within the widely held firm. It is hoped that, by identifying and responding to the dangerous coalition of financial interests at the

heart of the corporate governance process, the law might effectively counteract the expansion of market-based governance structures on an international level. This is an especially urgent task given the acknowledged informational weaknesses of the Anglo-Saxon model.

## Notes

1 For more details on the contractarian approach to the firm, see Weinstein, Gindis, and Manfrin, all in this volume.
2 We do not intend to decide who is the primary beneficiary of finance-led capitalism – shareholders or managers. For the purposes of our argument, it is sufficient to note that *both* are beneficiaries.
3 Note that Enron's board was composed of twelve "independent" directors out of a total of 14.
4 The conflicts of interest that used to run through the audit profession are put forward as the main factor explaining the corporate scandals of the Enron era (see, for example, Coffee 2002).
5 Here, the difference between the board of directors and the supervisory board is of little importance. In both cases, we are dealing with the central strategic organ of the firm.
6 For Zingales (1998) and for Blair and Stout (1999), though, rights of control over the board of directors should be vested in shareholders. One may be surprised by this conclusion, which seems to run counter to the idea of the board as an independent and neutral body, born out of an agreement amongst the various stakeholders. Blair and Stout (1999: 324), however, justify this allocation by pointing to the synthetic character of shareholders' interests: "Share value can sometimes be a proxy for, or an indicator of, the total value of rents being generated by the corporation. Not a perfect proxy, we believe, but at least it is one legitimate indicator." This is a classical, yet rather strange argument. Indeed, Blair (1995) had, in an earlier (and influential) work, defended the idea that the maximisation of stock price is different from the maximisation of the organisational quasi-rent as soon as workers invest in firm-specific human capital. Here, we find a mode of reasoning analogous to that of Williamson who, after recognising the importance of employees in terms of the creation of value (through their specific investments), excludes them from board-level participation.
7 On this point, see especially Berle (1965, reprinted in this volume).
8 This figure includes the District of Columbia. Avi-Yonah and Sivan (this volume) speak of this hostile takeover crisis in American corporate affairs as "a failed transformation."
9 The anomalous eight states where there is no such provision in force are: Alabama, Alaska, Arkansas, California, District of Columbia, Montana, New Hampshire and West Virginia.
10 Likewise, the Laws of New York make explicit reference in this regard to "the ability of the corporation to provide, as a going concern, goods, services, employment opportunities and employment benefits and otherwise contribute to the communities in which it does business": § 717(b)(v). The corresponding provision in force in the State of Missouri even goes so far as to adopt accounting terminology. In particular, it sanctions awareness by management to "[t]he future value of the corporation over a period of years as an independent entity discounted to current value," together with any "existing political, economic and other factors bearing on security prices generally or the current market value of the corporation's securities in particular": § 351.347. 1(1)(c), (2). The latter provision is further notable in so far as it highlights, as another managerial decision-making criterion, "[t]he financial condition and

earnings prospects of the person making the acquisition proposal including the person's ability to service its debt and other existing or likely financial obligations": § 351.347. 1(5).

11 British company law, in contrast, has traditionally avoided engagement with this underlying politico-economic conundrum, by virtue of the preclusive "proper purpose" doctrine applied by the English courts in regulating (*inter alia*) the conduct of a target company's management when faced with an actual or potential takeover bid. In essence, this doctrine serves to protect the accepted proprietary entitlement of shareholders, vested in them under the company's constitution, to consider the merits of any proposed bid personally, and to subsequently vote upon that proposal with a view to their own financial self-interest. To this end, defensive tactics by target boards are, as a general rule, forbidden under the common law, a position that is today backed up by the even stricter prohibition on "poison pill" strategies laid down by Regulation 7 and rule 21 of the City Code on Takeovers and Mergers. On the effect of the proper purpose doctrine and above City Code provisions generally, see Parkinson (1993): 140–46.

## Bibliography

Aglietta, M. and Rebérioux, A. (2005), *Corporate Governance Adrift: A Critique of Shareholder Value*, Cheltenham: Edward Elgar.

Alchian, A.A. (1969), "Corporate Management and Property Rights," in H. Manne (ed.), *Economic Policy and the Regulation of Corporate Securities*, Washington D.C.: American Enterprise Institute; reprinted in E. Furubotn and S. Pejovic (eds) (1974), *The Economics of Property Rights*, Cambridge: Ballinger.

Berle, A.A. (1932), "For Whom Are Corporate Managers trustees? A Note," *Harvard Law Review* 45(8): 1365–72.

—— (1947), "The Theory of Enterprise Entity," *Columbia Law Review* 47(3): 343–58; reprinted in this volume.

Berle, A.A. and Means, G. (1932), *The Modern Corporation and Private Property*; revised edn 1967, New York: Harcourt, Brace & World.

Biondi, Y. (2004), *La valorisation des actifs dans le cadre conceptuel de la future normalisation comptable international, particulièrement au regard des normes 36 et 38* [Recognition and measurement of assets according to the conceptual framework of coming international accounting harmonisation, especially the standards no 36 and 38], *Comptabilité Contrôle Audit*, 10 (2), December: 55–72.

Blair, M. (1995), *Ownership and Control. Rethinking Corporate Governance for the Twenty-first Century*, Washington D.C.: Brookings Institution.

Blair, M. and Stout, L. (1999), "A Team Production Theory of Corporate Law," *Virginia Law Review*, 85: 247–328.

Bratton, W. (2005), "The Academic Tournament over Executive Compensation," *California Law Review* 93(5): 000–000.

Cioffi, J. and Cohen, S. (2000), "The State, Law and Corporate Governance: the Advantage of Forwardness," in S. Cohen and B. Gavin (eds), *Corporate Governance and Globalization. Long Range Planning Issues*, Cheltenham and Northampton: Edward Elgar.

Coates, J.C. (1989), "State Takeover Statutes and Corporate Theory: The Revival of an Old Debate," *NYU Law Review* 63: 806.

Coffee, J.C. (1988), "Shareholders Versus Managers: The Strain in the Corporate Web," in J.C. Coffee, L. Lowenstein and S. Rose-Ackerman, *Knights, Raiders and Targets: The Impact of the Hostile Takeover*, Oxford: Oxford University Press.

—— (2002), "Understanding Enron: It's About the Gatekeepers, Stupid," *Columbia Law School Working Paper*, 207.

Cooper, S., Crowther, D., Davies, M. and Davis, E. (2000) "The Adoption of Value-Based Management in Large UK Companies: A Case for Diffusion Theory," paper presented at the 23rd annual Congress of the European Accounting Association, Munich, March 29–31.

Davies, P.L. (2003), "Workers on the Board of the European Company?," *Industrial Law Journal* 32: 75.

Deakin, S. and Morris, G.S. (2005), *Labour Law* (4th edn), Oxford: Hart.

Easterbrook, F.H. and Fischel, D.R. (1991), *The Economic Structure of Corporate Law*, Cambridge: Harvard.

Fama, E.F. (1980), "Agency Problems and the Theory of the Firm," *Journal of Political Economy* 88: 288.

Favereau, O. (1997) "*L'incomplétude n'est pas le problème, c'est la solution*," in B. Reynaud (ed.), *Les limites de la rationalité. Tome 2: les figures du collectif*, Paris: La Découverte.

Finkelstein, S. and Mooney, A.C. (2003), "Not the Usual Suspects: How to use Board Process to Make Boards Better," *Academy of Management Executive*, 17: 101–13.

Galbraith, J.K. (1973), *Economics and the Public Purpose*, London: Andre Deutsch.

Grossman, S. and Hart, O. (1986) "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy* 94: 691–719.

Hansmann, H. and Kraakman, R. (2002), "Toward a Single Model of Corporate Law?," in J.A. McCahery, P. Moerland, T. Raaijmakers and L. Renneboog, *Corporate Governance Regimes: Convergence and Diversity*, Oxford: Oxford University Press.

Holmström, B. and Kaplan, S. (2003), "The State of U.S. Corporate Governance: What"s Right and What's Wrong," ECGI Working Paper Series in Finance, 23.

Jensen, M.C. and Meckling, W. (1976), "Theory of the Firm: Managerial Behaviour, Agency Costs and Ownership Structure," *Journal of Financial Economics* 3: 305.

Jensen, M. and Murphy, K. (2004) "Remuneration: Where We've been, How We Got to Here, What are the Problems, and How to Fix Them," ECGI Working Paper Series in Finance, 44.

Kahn-Freund, O. (1956), "Legal framework," in A. Flanders and H. Clegg (eds), *The System of Industrial Relations in Britain*, Oxford: Blackwell.

Lichtenstein, N. (2002), "A Marx for the Shareholder Class," paper presented at the Gotham Center Conference: "New York City and the New Deal," New York, November.

Manne, H.G. (1965), "Mergers and the Market for Corporate Control," *Journal of Political Economy* 73: 110.

Millon, D. (1990), "Theories of the Corporation," *Duke Law Journal* April: 201–62.

Nader, R., Green, M. & Seligman, J. (eds) (1976), *Taming the Giant Corporation*, New York: Norton.

Parkinson, J. (1993), *Corporate Power and Responsibility: Issues in the Theory of Company Law*, Oxford: Clarendon Press.

Rebérioux, A. (2002), "European Style of Corporate Governance at the Crossroads: The Role of Worker Involvement," *Journal of Common Market Studies* 40: 111–34.

Robé, J.-P. (1999) *L'entreprise et le droit*, Paris: Presses Universitaires de France.

Roberts, J., McNulty, T. and Stiles, P. (2005) "Beyond Agency Conceptions of the Work of the Non-Executive Director: Creating Accountability in the Boardroom," *British Journal of Management* 16: 5–26.

Romano, R. (1996) "Corporate Law and Corporate Governance," *Industrial and Corporate Change* 5: 277–339.

Stewart, G.B. (1991), *The Quest for Value: The EVAī Management Guide*, New York: HarperBusiness.

Stone, C.D. (1975), *Where the Law Ends: The Social Control of Corporate Behaviour*, New York: Harper & Row.

Teubner, G. (1985), "Corporate Fiduciaries and Their Beneficiaries: A Functional Approach to the Legal Institutionalization of Corporate Responsibility," in K.J. Hopt and G. Teubner (eds), *Corporate Governance and Directors' Liabilities: Legal, Economic and Sociological Analyses on Corporate Social Responsibility*, Berlin: De Gruyter, 149–77.

—— (1994), "Company Interest: The Public Interest of the Enterprise 'in Itself'," in R. Rogowski and T. Wilthagen (eds), *Reflexive Labour Law: Studies in Industrial Relations and Employment Regulation*, Deventer: Kluwer.

Williamson, O. (1975) *Markets and Hierarchies: Analysis and Antitrust Implications*, New York: Free Press.

—— (1985) *The Economic Institutions of Capitalism*, New York: Free Press.

Williamson, O. and Bercovitz, J. (1996) "The Modern Corporation as an Efficiency Instrument: the Comparative Contracting Perspective," in C. Kaysen, (ed.), *The American Corporation Today*, Oxford: Oxford University Press.

Zingales, L. (1998) "Corporate Governance," in P. Newman (ed.), *The New Palgrave Dictionary of Economics and the Law*, London: Stockton Press.

—— (2000) "In Search of New Foundations," *The Journal of Finance* 55: 1623–53.

# Index

accountability, problem of 135–36
accounting: against black-box firm 249–251, 260n8; bank credit, business cycle and 84–85; bounded rationality 254; constituents 241, 244, 258; contractarian view of firm 239–40, 246–47; cost accounting 85–91; dynamics 246–47, 258; equilibrium theory 237–39; financial accounting 82–85; full process models 75–76; game theory 79–80; general equilibrium theory 75, 78–79; ideology and legitimation 76; incomplete contracts economics 239–40; macroeconomic accounting 73, 74; marginal cost pricing 238; mathematical economics 75, 79–80; measurement difficulties 74; microeconomic accounting 74–75; money, role of 78; moral hazard 243; national income 73; off-contracting 246; opportunity cost 86–89; records as economic statistics 85, 90–91; role playing 259; system and calculations 89–91; theory of firm and 82–91; transaction costs 88–89, 242–43; undertakers 259; valuation, theory of physical 148; value maximization, principle of 256
Accounting Research Association 83, 85, 87
accounting system: alternative to marginal cost pricing 241–42; dynamic function of 255–56; governance and disclosure 255–56, 257–58; institutional structure of production 245–52; understanding of firm as entity 248–52
agency theory 24, 32, 33, 242, 243; managerialism and 349–52, 370; shareholders and 324–25, 328
Aghion, P. and Tirole, J. 29
Aglietta, M. and Rebérioux, A. 33, 352
AIDS 153, 155, 174
Alchian, A.A. and Demsetz, H. 25, 34–35, 36, 37, 38, 267, 298
Alchian, A.A. and Woodward, S. 267, 268
Alchian, Armen A. 33, 267, 268, 277, 351
Allais, M. 327
altruism 71n4; in organisations 63
American Finance Association (AFA) 320, 321
American institutionalism 4–5, 6, 14
Anderson v. Abbott (1944) 187, 192, 195
Angell, J.K. and Ames, S. 157, 177n13
Anglo-Saxon concept of firm 366, 367, 369, 371
Annan, Kofi 153
Anthony, Robert N. 11, 15, 17, 201–15, 249, 256, 308–9, 319, 328
anti-takeover legislation 364–66
Antonelli, Etienne 150n8
Aoki, M. 21, 38, 39
Archer, Margaret S. 281
Arrow, Kenneth J. 254
Arrow, K.J. and Debreu, G. 78
Arthur, Brian 327
artificial entities 158–59, 162, 163, 169–70, 186
authority: control and 26–27; coordination and 67; effective authority 29; and employment relation 59–61; hierarchy and 25, 27, 39; profit as revenue of 136–37

Avi-Yonah, Reuven S. 9, 14–15, 153–85, 272, 294, 308, xi
Azzini, L. 122

Babbage, Charles 38
Bacon, Francis 119
Baker, George P. 237, 246, 251, 257
Baker, G.P. et al 29
Bank of Augusta v. Earle (1839) 160
Bank of US v. Dandridge (1827) 157, 159
Bank of US v. Deveaux (1809) 157–58, 159, 160, 161
Barzel, Y. 33
Baudry, B. and Gindis, D. 285
Baumol, W. 109, 112, 124
Baxter, W.T. 87
Bazzoli, L. and Kirat, T. 150n18
Bebchuk, L.A. and Fried, J. 176–77n4
Beigel, R. 298
benchmarking 355–57
Bentham, Jeremy 107
Berle, A.A. and Means, G.C. 10, 14, 17, 23, 46, 307, 348–50, 357–60, 362–64; on firm and ownership 32–33
Berle-Dodd debate 154, 176n2, 176n4, 179n33, 273, 307
Berle, Adolf A. Jr. 4, 5, 10, 13, 15, 92–104, 112, 131, 138, 186–200, 274, 276, 284, 293, 294, 301, 307
Besta, Fabio 111, 119, 120, 126, 127
*Betriebswirtschaftslehre* 114–16, 117, 118, 121, 122, 123, 124, 126, 128; economic transformation, firm as system of 122–23; firms and markets according to realism 121–22; freedom of action 114–15; income-producing combinations 123–24; profitability coordinations 123–24; spirit in men 115
Bettis, R.A. 334
Beveridge V. N.Y. Elevated R.R. (1889) 165
Biondi, Yuri 3–8, 9–18, 123, 132–33, 148, 237–65, 281–82, 295, 297, 299, 305, 319, 329, 336, 340, 352, 356, xi
Black, B. and Kraakman, R. 176n4
Black, F. and Scholes, M. 322, 323
Blackstone, W. 156–57, 177n12
Blair, Margaret M. 160, 272, 273–74, 302
Blair, M.M. and Stout, L.A. 25, 26, 32, 36–38, 175–76, 360–62, 371n6
Bline, D.M. and Cullinan, C.P. 336
Blumberg, Phillip I. 160, 172, 284
Böhm-Bawerk, Eugen von 107, 111, 127
Boisguillebert, Pierre le Pesant, Sieur de 73
Bolton, P. and Scharstein, D.S. 25
Boulding, K.E. 117
Boutroux, Émile 114
Bowles, S. 37
Bradley, A.C. 113
Branson, D.M. 154
Bratton, William W. 179n34, 270, 286n3, 354
Brennan, M.J. 322
Britain 103, 113, 205, 368, 369; British company law 372n11
Brodbeck, May 279
Broglia, G. 122
Buchanan, J.M. 87
Bunge, Mario 276, 278
Burrill v. Nahant Bank (1840) 178n16
Bush, President G.W. 153
business studies, development of 111
Butler, H.N. and Ribstein, L.E. 273

Cannan, E. 111
Cantillon, Richard 107
Canziani, A. and Rondo-Brovetto, P. 114, 129n2
Canziani, Arnaldo 3–8, 9, 14, 107–30, 132–33, 148, 294, xi
capital: Capital Asset Pricing Model (CAPM) 324, 333–34; capital budgeting 208; capital sources 95–96; control over application of 92; marginal efficiency of 108
capital structure: decisions on 336; dynamic satisficing 340–41; entity-firm perspective 339–42; irrelevance of 319; rebalancing dynamically through time 341–42
capitalism: capitalist control, shift from 96–97; enterprise and 133, 148–49; firm as main institution of 21, 22, 45; IFRS and 132; industrial and financial 3–4; legal concepts and 138–39; and private enterprise 93
Cassady Jr, Ralph 97–98
Cassel, Karl Gustav 127
causal power: casual power and real entity 278–79; casual reciprocity 279; efficient and final causation 279
central planning 58, 67–68
Cerboni, Giusseppe 119, 126, 127
Chandler, Alfred D. 10, 21, 22, 29, 39–40, 43, 45, 284
Chen, R. and Hanson, J. 154, 175

Cheung, Steven N.S. 266, 267
China 56; society in 64
Chisholm, Roderick M. 277
Cintas v. American Car and Foundry Co (1942) 195
Cioffi, J. and Cohen, S. 352
Clark, J.B. 111, 217, 220
Clark, John Maurice 4, 14, 132; enterprise as object of social control 141–44
Clark, R.C. 156
classical economic theory 97–101, 104n6; capital 92, 93, 95–96, 96–97; competition 92, 97–98; corporate conscience and power 99–101; economic control, elements of 99–100; firm, impact on 92–104; long-range planning 99; management responsibility and 92–93; owner-entrepreneurship 94, 101–2; power over capital 95–96; profit maximization 98–101; risk-taking 101–2; shareholder responsibility 96; social-political factors, stockholders and 101–3; stock and securities markets 92, 101–3; tradition and 93; wealth holding 94–95
Clinton, President W.J. 153
Coase, Ronald H. 7n1, 9, 10, 12, 13, 22, 82–91, 237, 239–40, 246, 254, 257, 268, 298; legacy of 23–32, 125–26
Coates, J.C. 365
Coffee, J.C. 366
cognitive coordination 42, 43
cohesiveness 279–80
common interest 306–9
Commons, John Roger 4, 11, 14, 46, 132, 149, 246, 280, 295, 303; legal-economic dynamics of enterprise 144–48; original perspective on enterprise 145–46
company interest 364–66
competition 97–98; price system 5; process 92
composite entities 195–96
conflict of interests 307; and legitimacy of private power 142–44
constituents: in accounting 241, 244, 258; common interest of constituencies 306–9; constitutive rules 45–46; legal relations between 302
continuity, emphasis on 339–42
contractual perspective: contractarian view in accounting 239–40, 246–47; coalition 35–36, 176, 267–69, 293, 313n34, 337–39, 367; contractual relationships over time 27–28; enterprise entity, firm as 267–68; firm, economic theory of 23–32; governance 359–62; holistic competence perspective and 283–84; law 298–300, 308
control: authority and 26–27; capital, control over application of 92; capitalist control, shift from 96–97; control undertakings, functionaries and 225–26; cost control 117; economic control, elements of 99–100; enterprise as object of social control 141–44; intra-association control 276; market control 99; non-negotiated rules and social control 141–42; primacy of 222, 223; residual control rights 26, 27, 28; types of 221–22, 224; ultimate control 219; *see also* ownership
conventional wisdom, reactions against 108–9, 109–12
convergence of interest 352–57, 370–71
Cook, W.W. 165, 166
Cooper, S. *et al* 353
coordination: and control 99–100, 242, 252–53, 254–55, 303–4; establishment in law 303–4; between and within organisations 68–69; in organisations 66–69; between people 24; between workers 27
Copeland, Morris A. 277
Copeland, T.E. 332
Copp, David 279, 280, 281
Coriat, B. and Weinstein, O. 42
corporate accountability 358–59; decline in 354
corporate social responsibility (CSR) 153–54, 162–66, 173–76; aggregate view 157–64, 168, 172–73, 178n21; artificial entity view 158–59, 162–63, 169–70, 186; closely-held enterprises 155, 161; corporate personality 166–67; corporation as separate entity 154–56; cyclical transformations of form 155–73; diversity jurisdiction 157–58; general incorporation laws 159–60; historical perspective 154–73; hostile takeover crisis 166–71; limited liability, rise of 159–60; management power 175–76; membership corporation 155, 156–58; multinationals 155, 171–73; non-profit corporations 155, 156–57; private corporations 155,

158–59, 161–66; public corporations 155, 161–66; real entity view 159–60, 162–64, 166–68, 174–76; shareholder aspirations 154, 171, 174–75; social goals 175–76; *ultra vires* doctrine 164, 165, 166, 179n28
cost accounting 85–91; profit maximization and 207–8
costs: capital structure decisions and values 336; financing values or 339–40; revenues and values 330–31
Cournot, A.A. 78, 79, 107
Covaleski, M.A. 148
Croce, Benedetto 113, 114, 128
CTS Corp v. Dynamics Corp of America (1987) 169–70
Cutler, D.M. *et al* 327
Cyert, R.M. 109
Cyert, R.M. and March, J.G. 56

Dartmouth College Trustees v. Woodward (1819) 157, 158–59
Davenport, H.J. 111
David, Donald K. 202–3
Davies, P.L. 368–69
De La Vergne Co v. German Sav Inst (1899) 166
De Long, J.B. *et al* 327
Deakin, S. and Morris, G.S. 368
Deakin, Simon 284
Dearborn, D.G. and Simon, H.A. 65
debentures 196–97
Debreu, Gerard 74, 77, 78
decision-making: entrepreneurial activity 220–21, 224, 225; organisations 65; power, ownership and 33–34
Deiser, G.F. 178n21
Dejnožka, Jan 275, 277, 278, 279, 282
Demaria, G. 112, 302–3
Demsetz, H. and Lehn, K. 58, 66
Demsetz, Harold 33, 243, 268
Denrell, J. *et al* 330
Descartes, René 119
Devons, Ely 67
Dewey, John 113, 166
Dietrich, Rudolf 118
Dobb, Maurice 232n15
docility 71n4; in organisations 64
Dodd, E. Merrick 179n33, 273, 307
Dodge v. Ford Motor Co (1919) 176n4
Donaldson, T. and Preston L.E. 273
Dosi, G. and Coriat, B. 44
Dosi, G. and Marengo, L. 40
Dosi, G. *et al* 41–42, 43
Drucker, Peter 110
Dubey, P. and Shubik, M. 79
Duff, P.W. 156
Dufrene, U. and Wong, A. 332
Dugger, W.M. 145–46
Dupuit, Arsène-Jules-Étienne-Juvénal 107
Durand, P. 140
Durkheim, Emile 10–11
Dwight, T.W. 165
Dybvig, P. and Ross, S.A. 322
dynamic efficiency 28, 41
dynamic entity, firm as 6–7, 9, 16–18, 21, 41
dynamic function of accounting system 241, 245–52, 255–56, 258

Easterbrook, F.H. and Fischel, D.R. 168, 174–75, 268, 273, 284, 351, 352
*Economia aziendale* (EA) 119–21; accounting system, integration of 120, 121; as autonomous discipline 126–27; ‘critical positivism’ 119–20; firms and markets according to realism 121–22; income-producing combinations 123–24; innovative concept of firm 120, 122–23; profitability coordinations 123–24
economics: ‘anti-realistic’ firm 107–9; ‘black box’ perspective in 5, 6, 13, 109, 125, 131–32, 241–42, 260; business studies, development of 111; conventional wisdom in 109–12; dynamic and enterprise 134–35; Economic Value Added (EVA) 354–55, 366; enterprise as institution 133–38; firm as non-theoretical object 107–9; going concern, enterprise as 144–46; high theory in 4, 9, 10, 13, 14; innovation, fervour for 110–11; marginalism 108; neglect in history of 127–28; neo-classical tradition 3–4; new institutional economics 55–56; parasitism 143; price discovery 125–26; psychological approach 108–9; public order and private order 146; revenue 135–36; rules and enterprise 146–47; social sciences, methodolodies of 112–13; substance of firm 292–96, 311n1; worldview in 3, 5
Economic Value Added (EVA) 262n46, 354–355, 366 *see also* shareholder value
Edgar v. MITE Corp (1982) 167–69, 171
Edgeworth, Francis Ysidro 109
Edison, Thomas A. 126

Edwards, Ronald 83, 85–87
Eisenberg, Melvin A. 270
Eisenhower, President D.D. 211
Eliasson, Gunnar 283–84
Ellerman, D.P. 27
Ellsberg, D. 327
Emmeche, Claus *et al* 278
employment relation 28–29, 30, 35; authority and 59–61
enterprise as institution 133–38
enterprise entity, firm as 9–10, 15–16, 294–95, 358; aggregates and fictions 271–76; causal power and real entity 278–79; cohesiveness 279–80; composite entities 190–92, 195–96; contractual perspective 267–68, 283–84; '*de facto*' corporation 188–90; debentures 196–97; defective incorporation 188–90; durability through time 279–80; economic coordination 252–55; enterprise risk 284–85; finance and economics 17, 337–39; governance 348–49, 362–70; heterogeneity 283–84; holism and emergent properties 278–79, 283–84; institutional perspective 248, 258, 280–82; interactive complex systems 282, 285; law and economics 16–17, 131–50, 292–314; legal entity 188–93; mereological essentialism 277; multi-personality 191–92; multi-unit structures 284–85; novelty and emergence 278–79; objectionability of enterprise 193–95; ontological individualism 279–80, 282–83; partnership and incorporation 190–91; personality 186–87, 274–75; real entity theory 274–76, 285; reductionism 277–79; representation principle 275–76; special liability enterprises 192–93; subsidiary agreements 196–97; transactional perspective 248, 258; whole and dynamic system 245–48, 257–58; *see also* firm, economic theory of
enterprise risk 284–85
entrepreneurial activity: control primacy 222, 223; control types 221–22, 224; decision-making 220–21, 224, 225; definition of firm, problem of 228–31; discontinuous innovation 224; entrepreneur and new theories of the firm 243–45, entrepreneur and legal and accounting facts 243–44; firm as entrepreneur 216–33; functionaries and 218–19, 225–27; government responsibilities 218, 222–23; historical concept of 216–17; labour and 220, 224–25; ownership and risk 218; profit concepts and incomes 225, 230–31; proprietary capitalist management 217; provisional guaranty 218, 223; remuneration 225–27; risk-bearing and management 217–19, 223, 228; ultimate control 219; uncertainty and 230–31; working rules, collective will and 229
equilibrium economics 11–12, 237–39; *see also* general equilibrium model
equilibrium values and risk 333–35
ethics: ethical-political view of enterprise 16–17; problem of profit maximization 211
Europe 3, 5, 56; continental accounting and business economics 4, 5–6, 14; dissatisfaction with economics in 109–12; ICE Directive 367; liberal political philosophy 24; methodological controversy in 110; scientific methodology in 112–14; 'social' model 366–70; Societas Europaea (SE) 368–69; theories of the firm in 128; Works Council (EWC) Directive 367
evolutionary perspective 38–44
expectations 253; homogeneity of 330

Fama, E.F. and Miller, M. 324
Fama, Eugene F. 34, 267, 325, 351, 359
Favereau, Olivier 246, 362
Ferrara, Francesco 111
Fetter, Frank Albert 111
fictions and aggregates 272–74
fiduciary responsibilities 361, 366
finance and economics 317–44; assumption and reality 319–21; continuity, emphasis on 339–42; costs, financing values or costs 339–40; enterprise entity, firm as 17; entity-firm finance 324, 337–39; finance theory and the firm 323–25; history of finance theory 318–23; methodological individualism 319; methodology 318–23; Net Present Value (NPV) 325, 334, 336, 338, 342; predictive effectiveness 319; shareholder wealth maximization (SWM) 324, 325, 329, 331–33, 337, 344n17; slacks 337–39; traditional finance 317–23; uncertainty and continuity 337–38; unsystematic

risk 333–35; values or costs 339–40; *see also* neo-classical finance theory (NFT)
financial accounting 82–85
Findlay, M.C. and Williams, E.E. 324, 334
Finkelstein, S. and Mooney, A.C. 353
firm, economic theory of: collective skills and organisational learning 42–43; competence-based perspective 10, 23, 31, 40–44, 49n43, 282–83; constitutive rules 45–46; contractual incompleteness 10–11, 23–32, 24, 25–26; dynamic entity, firm as 6–7, 9, 16–18, 21, 28, 41; employment relationship 28–30, 35; evolutionary perspective 38–44; firm and ontological glue 266, 277–79, 282; firm and stickiness 280, 286n9; firm as becoming concern 252–53; firm as institution 45–47; horizontal coordination 36; knowledge-based perspective 23, 31, 40–44, 49n43, 282–83; legitimacy 29, 44; neo-institutional perspective 23–32; organisational architectures 21, 38–44; ownership logic against institution logic 359; partnership model 36; political coordination 42–44; power and property 30; production capabilities 29–30, 38–44; public corporation 32–38; residual control rights 26, 27, 28; residual income 35–36; resource-based perspective 40; risk sharing 35–36; shareholder-manager relations 23, 32–38; skills and organisational learning, social nature of 42–43; social organisation 32–33; stakeholders 35, 36, 47; team approach 25, 36–39; transaction costs theory (TCT) 24, 25–26; *see also* authority; enterprise entity, firm as; ownership; property rights
Fischel, D.R. 154, 168
Fisher, Irving 111, 119
Fowler, Ronald 83, 86
Fox, M.B. 172
Frankfurter, G.M. and McGoun, E.G. 326, 327
freedom of action 114–15
freedom under law 292–93
French, Peter A. 279, 281
Freund, Ernst 17, 178n21, 274, 275–76, 278, 281
Friedman, Milton 15, 154, 175, 319, 322, 326
Friend, I. 321
Galbraith, John Kenneth 350, 356
game theory 79–80; in organisations 66
gatekeepers 351–52, 356
General Accepted Accounting Principles (GAAP) 261n28
general equilibrium model 75; importance of 79–80; limitations of 76–77, 78–79; nonstrategic nature of 78; price system, role of 78–79; static nature of 77–78; timeless nature of 78
General Motors 57, 65, 71n5, 94, 101–2
Gentile, Giovanni 113
Germany 3, 5, 14, 111, 369; Nazi Germany 103, 107; scholarship in, range of opinions of 117–19
Giannessi, E. 121
Gibbons, R. and Murphy, K.J. 237, 246, 251, 257
Gibbons, Robert 24–25, 27, 32, 283
Gierke, Otto Friedrich von 274
Gindis, David 10, 16–17, 266–91, 297, 304, xi
Goldstein, Jeffrey 278
Gomberg, Moses 126
Gordon, Cosmo 83
Gossen, H.H. 107
Gould, J.R. 87–88, 90
governance 348–71; agency theory, managerialism and 349–52, 370; anti-takeover legislation 364–66; benchmarking 355–57; company interest 364–66; contractarian approach 359–62; convergence of interest 352–57, 370–71; corporate accountability 358–59; corporate politicization 348; Economic Value Added (EVA) 354–55, 366; fiduciary responsibilities 361, 366; gatekeepers 351–52, 356; government responsibilities 218, 222–23; information exchange 351–52; institutionalisation 357–59; managerialism, agency theory and 349–52, 370; non-shareholder constituents 366–70; normative assessment 357–62; positive assessment 349–57; pro-shareholder mechanisms 348–49; quiescent shareholders 350; Return on Equity (ROE) 355; shareholder primacy 351–52, 353, 358, 361–62; shareholder value 353, 356, 370; stock-price mechanism 351–52; technostructure 350–51; value based management 353, 354

Graham, E.M. and Krugman, P.R. 172
Grandmont, L.M. and Younes, Y. 78
Green, L.R. 78
Greenwood, Daniel J.H. 176–77n4, 273, 277
Grossman, S. and Hart, O.D. 269, 360
Grossman, S. *et al* 26
groups: and companies, juridical differentiation 305; industrial group as unitary enterprise 304–6; relations between 24
Gruchy, G. Allen 4, 5
Gynther, S.R. 301

Hale v. Henkel (1906) 162–64
Hansen, Alvin Harvey 127
Hansen, C. 167
Hansmann, H. and Kraakman, R. 160, 176n4, 369
Hansmann, Henry 268, 273
Harbrecht, Dr Paul 102
Harris, R. 171
Hart, O.D. and Moore, J. 27, 131, 266, 269, 300
Hart, Oliver D. 27, 30, 269, 271, 272, 300
Hauriou, Maurice 150n10
Hawkins-Simon theorem 68
Hawley, Frederick B. 217
Hayek, F.A. von 48n21, 67, 68, 86
Hegel, Georg W.F. 113, 115, 116
Heisenberg, Werner 212
Hepworth, S.R. 117
Hicks, J.R. 127
hierarchy and authority 25, 27, 39
high theory in economics 4, 9, 10, 13, 14
Hilferding, R. 110
history: of corporate social responsibility (CSR) 154–73; of entrepreneurism 216–17; of finance theory 318–23
Hodgson, G.M. and Knudsen, T. 280, 284
Hodgson, Goeffrey M. 22, 43, 46, 243, 275, 278, 280, 281, 284
Hohfeld, Wesley 144
Holmes, O.W. 173
Hölmstrom, B. 26, 27, 29–30, 32, 243
Hölmstrom, B. and Kaplan, S. 353
Hölmstrom, B. and Roberts, R. 29
*homo oeconomicus* 323, 327
Hoover, Edgar M. 214n5
Horwitz, M.J. 161
hostile takeover crisis 166–71
Hoyt v. Thompson's Executor (1859) 178n16
Husserl, Edmund 113, 118, 128, 129n3

identity and enterprise 281–82
ideology and legitimation 76
IFRS (International Financial Reporting Standards) 132
Ijiri, Y. and Simon, H.A. 57, 69
Ijiri, Yuji 79, 81n4, 249, 260n8
income: distribution, inequalities in 5; of enterprise, institutional question of 136–37; of firm and of owners 244–45; importance of 120; income-producing combinations 123–24; net income of the firm 256; streams in time, relevance of 125
incompleteness; incomplete contracts 10–11, 26–27, 31, 55, 59, 144, 239–40, 242–245, 299, 312n18, 348, 360–62
industrial evolution 93
information: decentralisation of 38–40; free availability of 321, 322; information exchange 351–52
innovation: fervour for 110–11; innovative concept of firm 120
institutions: development in classical economic theory 93; enterprise as institution 147; individuals and 280–82; ownership logic against institution logic 359; revenues of 135, 137; and social organisation 32–33; system of 46–47
integrated enterprise, theory of 305–6
interaction 245, 259; between agents, freedom in 24; interactive complex systems 282, 285
interdependence: and reward 61–62; of teams 37
interdisciplinarity 16–18
internal governance 309–11
intra-association control 276
investor rationality 321, 322
Ireland, Paddy 273
Italy 3, 5, 14, 111, 118; *Economia aziendale* (EA) 119–21, 122, 123, 124, 127, 128; schisma in 119–21
Iwai, Katsuhito 273

Jaeger, P.G. 307
James, Emile 150n11
Jensen, M.C. and Meckling, W. 33–35, 38, 266–68, 270–73, 275, 282, 322, 324–25, 351
Jensen, M.C. and Murphy, K. 353, 354
Jensen, M.C. and Smith, C.W. 318, 319, 326
Jensen, Michael C. 154, 332, 338, 340
Jevons, William Stanley 107

judiciary: common interest 310; partial nature of approaches 296–304

Kahn-Freund, O. 368
Kahneman, D. and Tversky, A. 327
Kahneman, D. *et al* 333
Kant, Immanuel 113, 115, 119, 128
Kay, Neil M. 280
Keasbey, E.Q. 166
Keynes, John Maynard 80, 108, 126, 127–28
Khalil, Elias L. 276
Kirat, T. and Bazzoli, L. 243
Kirat, Thierry 3–8, 9, 14, 131–52, 302, xi
Kirman, Alan 238, 260n8
Kirzner, Israel 294
Klein, B. *et al* 25
Knight, Frank H. 15, 217, 219–20, 221, 224, 225, 231–32n11
knowledge: knowledge-based perspective on firms 10, 23, 31, 40–44, 49n43, 282–83; and learning, central position of 40–44
Koopmans, T.C. 75
Kornai, J. 56
Krafcik, J. and Wokack, J.P. 65
Kuh, Edwin 208
Kuhn, Thomas 113

Langlois, R.N. and Foss, N.J. 280
Larcom, R.C. 166
Laski, Harold J. 178n21, 274
law: artificial person, firm as 296–97; common interest of constituencies in firm 306–9; conflict of interests 142–44, 307; contractual perspective 298–300, 308; corporation as enterprise entity 187; economic coordination 303–4; economic substance of firm 292–96, 311n1; economics and law, synthesis between 293; egoistic interests 308; enterprise and 139–40, 147; external regulation 309–11; firm as co-law maker 303–4; firm as entity 292–314; firm as social organisation 295–96; freedom under law 292–93; industrial group as unitary enterprise 304–6; judicial approaches 296–304; legal-economic framework 14, 16–17, 295–96, 299; legislatures and contractual relations 143–44; legitimacy and structure 44; liability in unitary firms 306; limited liability companies 301, 310; limited partnerships 292; nexus of contracts, firm as 298–300; private and public order 146; private power, legitimacy of 142–44; registration 292–93, 294–96; regulation, scope of 292–93; rules and enterprise 146–47; social institutions 302–3; *see also* governance; property rights
Lawson, Tony 281
Leitner, Friedrich 118
Leontief input-output models 68
Leslie v. Lorillard (1888) 165
Levitt, Theodore 154
Licht, A.N. 172
Lichtenstein, N. 349
limited liability companies 310; characteristics of 301; rise of 159–60
limited partnerships 292
Lintner, J. 324, 333
liquidity principle 336, 338
Littleton, A. Charles 4, 14, 249, 251, 260n15, 261n27, 261n28, 262n44, 312n21, 313n37
London School of Economics (LSE) 82–83, 86, 87
Louisville Railroad Co v. Letson (1844) 160

McAndrews & Forbes Inc 164
Mach, Ernst 110, 114, 119, 122, 128
Machen, Arthur W. 166, 178n21, 274
Machlup, F. 109, 112, 124, 311n5, 331
MacMillan, K. and Farmer, D. 281
macroeconomic accounting 73, 74
Maitland, Frederick W. 274, 275
Mäki, Uskali 276
management: function, shift in 92; and income, relevance of 124–26; managerial accountability, decline in 354; managerialism, agency theory and 349–52, 370; ownership, management and control 243–44; performance, measurement and 175; power and real entity theory 175–76; profit maximization and management practice 208–9; proprietary capitalist management 217; responsibility and classical economic theory 93; risk-bearing and management 217–19, 223, 228; value based management 353, 354
Manfrin, Federico 10, 16, 17, 131, 284, 292–316, 358, xi
Manne, H.G. 351, 366

Manne, H.G. and Wallich, H.C. 176n4
March, J.G. 109
March, J.G. and Simon, H.A. 56
Marengo, L. and Dosi, G. 39, 40
Marengo, L. *et al* 39
marginal analysis, deployment of 204–5, 207
marginal cost pricing 238
Margolis, J.D. and Walsh, J.P. 153, 154
Margolis, Julius 214n5
markets: classical economic perspective on 54, 57; dominance and corporate social responsibility (CSR) 153–54; dualism between firm and market 240, 246; and firms, opposition between 25; importance for contemporary discourse 21; as ineffective regulatory system 309–10; market competitiveness, perfection in 321, 322; market control 99; market operation 88–90; market or organisational economy? 56–57; market transactions 56; neoclassical economic perspective on 54–55, 57; new institutional economics and 55–56; and organisations, boundary between 57–58; organisations and 11, 54–70; boundary between 57–58; pricing practices and market exchange 67–68; relative efficiency of organisations and 58
Markowitz, H. 318, 333
Marshak, J. and Radner, R. 38
Marshall, Alfred 108, 109, 111, 127, 128, 217, 231n4
Marshall, John 157–60
Marshall v. Baltimore & Ohio R.R. (1853) 178n15
Marx, Karl 27, 38, 107
Marzo, Giuseppe 10, 17, 317–47, 357, xi
Masini, C. 121–22
Mason, Edward S. 212
Mason v. Pewabic Mining Co (1890) 178n20
Masten, S.E. 24, 29
mathematical economics 75, 79–80
Mattei, U. 301, 302
Means, Gardiner C. 4, 92
membership corporation 155, 156–58
Menger, Carl 107
mereological essentialism 277
Merton, R.C. 322
Meyer, John R. 208
microeconomic accounting 74–75
Mill, John Stuart 107, 114, 119, 123, 126, 127–28
Miller, M. 323
Miller, M. and Modigliani, F. 318, 322, 323, 324, 340
Millon, D. 176–77n4, 365
Mises, Ludwig von 294
Mitchell, L.E. 176–77n4
money, role of 78, 256
Moore, John 47, 48n23, 131
Moore, Marc T. 286n7, 313n34
moral hazard 243
Morawetz, Victor 272
Mossin, J. 324, 333
motivation 54, 58–59
Muchlinski, P.T. 172
multi-unit structures 284–85
multinational enterprises 155, 171–73

Nader, R. *et al* 366
national income 73
Natorp, Paul 113
natural selection and fitness 63
Nelson, R.R. and Winter, S.G. 40
neo-classical finance theory 317–18, 319, 321–22, 323–24, 339, 340, 342, 343; bounded rationality 330, 337–38; capital structure, decisions on 336; core propositions 321; costs 330–31, 336; criticism of 325–36; dynamics, statistics *versus* 329; entity-firm concept 328–29; equilibrium, values and risk 333–35; expectations, homogeneity of 330; firm as production function 323–24; full rationality 330, 337–38; *homo economicus* and 323, 327; information, free availability of 321, 322; investor rationality 321, 322; law, firm and finance theory 335–36; liquidity principle 336, 338; market competitiveness, perfection in 321, 322; no-arbitrage assumption 321, 322–23; normative approaches 325–26; positive approaches 325–26; rationality 330, 337–38; revenues, values and costs 330–31; shareholder wealth maximization (SWM) 321, 322; shareholders and firm 330–33; statistics *versus* dynamics 329; system, complexity and 329; value-additivity 329; values and 326–27, 330–31, 333–36; wealth maximization 331–33
neo-institutional perspective 23–32, 299
Net Present Value (NPV) 325, 334, 336, 338, 342
Neumann, J. and Morgenstern, O. von 321

new institutional economics 55–56
Nicklisch, Heinrich 5, 14, 111, 113, 249; *Betriebswirtschaftslehre* (BWL) 114–16, 117, 118, 121, 122, 123, 124, 126, 128
non-profit membership corporations 155, 156–57
non-profit organisations 58
Nonaka, Ikujiro 41
Nonaka, Ikujiro *et al* 284
novelty and emergence 278–79

O'Connor, M.A. 176–77n4
off-contracting 246
Ohnism 44
Onida, P. 121
ontological commitment 270–71
ontological individualism 279–80, 282–83
opportunism 243, 245, 254; hold-up and 245, 246
opportunity cost 86–89
organisational architectures 9, 10–13, 21, 38–44
organisations: accounting statements 65; altruism 63; authority and coordination 67; authority and employment relation 59–61; central planning 58, 67–68; coordination 66–69; decision-making 65; divisionalization 62; docility 64; efficiency in 58–59; employment relation, authority and 59–61; free-riding within 62–63; game theory 66; interdependence and reward 61–62; loyalty and organisational goals 62–66; market or organisational economy? 56–57; motivation in 54, 58–59; natural selection and fitness 63; nonprofit organisations 58; organisational identifications 64, 65; ownership concentration 58; pricing practices 67–68; profit maximization 58–59, 60; responsibility for results 65; rewards as motivations 61–62; self-interest 58; simplification and focus 64–65; social organisation 58; survival traits 63–64; ubiquity of 56–58
Orts, Eric W. 167, 285
owner-entrepreneurship 101–2
ownership: concentration of 58; and control 23, 26–27; separation of 31, 32–33; 'dark side' of 43, 246, 258; depersonalization of 358; distinction between firm-entity and 243–44; distribution of 94; logic as opposed to institution 359; management and control 243–44; productive assets, ownership of 29–30; and resources 349–50; and risk 218

Pacioli, Fra' Luca 73, 111
Paine, C.L. 86
Pantaleoni, Maffeo 81n3, 127
Paramount Communications v. Time Inc (1989) 170–71
parasitism 143
Pareto, Vilfredo 107, 108, 109, 126, 260n20; Pareto distributions 57; Pareto optima 42, 71n6
Parkinson, J. 364, 366
partnership and incorporation 190–91
partnership model of firm 36
Paton, A.W. 111, 126, 233n27
Paton, A.W. and Littleton, A.C. 15, 231n4, 232n21
Pembina v. Pennsylvania (1888) 162
Penrose, E.T. 40, 124
Perroux, François 4, 5, 6, 12, 14, 45, 112, 132, 148–49, 238; economic perspective of 133–38
personality, notion of 229, 274–75
personification of the firm 243; 294
Pescatore, M. 310
Peterson, Shorey 93, 95, 97, 98, 100, 103
Pettit, Philip 279, 280
Petty, Sir William 73, 107
Pigou, Arthur Cecil 128
Plant, Arnold 83, 86
Poincaré, Henri 110, 114, 119, 122, 128
Polanyi, Karl 261n25
political coordination 42, 43, 44
political-economic factors in classical economics 100
Pound, Roscoe 144
Powell, Colin 153
power: over capital 95–96; private power, legitimacy of 142–44; and property 30
price: discovery in economics 125–26; pricing and market exchange 67–68; pricing for profit maximization 205–8; re-examination and profit maximization 205–7; supply and demand 88–90; system 239, 254
private corporations 155, 158–59, 161–66
private power, legitimacy of 142–44
*Privatwirtschaftslere* 114; cost control 117; exchange 117; experience as foundation of scientific knowledge 116; individual advantage and 116–17; relevance of income in 116–17

production: activity of 40; and capabilities 38–44; institutional structure of 12–13; productive assets, ownership of 29–30
profit: of business corporations 155, 157; concepts and incomes 230–31; income and entrepreneurial activity 225; motive in organisations 58; profitability coordinations 123–24; as revenue of authority 136–37
profit maximization 201–14; capital budgeting 208; classical economic theory 98–101; consistency with management practice 208–9; cost accounting and 207–8; courses constructed on premise of 203–4; difficulty of 210–11, 213; ethical problem of 211; goals, realism in 202–5; immorality of 210–11, 213; long-run maximization 208–9, 209–10; marginal analysis, deployment of 204–5, 207; most useful full cost, construction of 211–13; policies and practices 208–9, 210; pricing 205–8; problem of 15; satisfactory return 202–3, 204–5, 212, 213–14; social pressures 211–12
property: and classical economic theory 93; depersonalisation of 46; regime and form of organisation 29
property rights 239, 242, 243, 245, 254; firm as object of 300–303; firm-entity approach 302; shareholders and 301–2; theory of 24–25, 26, 27, 29, 33–35, 37
proprietary capitalist management 217
provisional guaranty 218, 223
public and private order 146
public corporations 32–38; corporate social responsibility (CSR) and 155, 161–66
Putterman, L. and Kroszner, R.S. 29

quiescent shareholders 350

Raby, William L. 262n44, 280, 312n24
Radner, R. 36, 38–39
Rajan, R.G. and Zingales, L. 30–31, 37, 38, 131, 237, 246, 257
Rathenau, Walter 4, 14, 33, 358, 359
rationality, bounded *versus* full 330, 337–38
real entity: as response to opponents of corporate social responsibility (CSR) 174–76; theory of: implications for enterprise of 285; revival of 274–76; view of corporations 15–16, 159–60, 162–63, 164, 166–67, 168
Rebérioux, Antoine 10, 17, 131, 293, 299, 311, 348–74, xii
records as economic statistics 85, 90–91
reductionism, failure of 277–79
registration in law 292–93, 294–96
regulation, scope of 17, 292–93
Reich, R.B. 172
Reichert, A.K. *et al* 154
relational firms 285
relationships 259; firm as system of 241; fundamental relationships 253
remuneration 225–27
Renouvier, Charles-Bernard 113
representation principle 275–76
residual control rights 26, 27, 28
residual income 35–36
Return on Equity (ROE) 355
revenues: theory of creation and distribution of 135–36; values and costs 330–31
Revlon v. MacAndrews & Forbes Inc (1986) 170
rewards as motivations 61–62
Ribstein, Larry E. 273, 286n2, 286n5
Ricardo, David 107, 108, 119
Ricci, U. 111
Richardson, G.B. 5
Rickert, Heinrich 113
Rieger, Wilhelm 118
rights of corporations, regulatory powers and 162–66
Ripert, Georges 14, 131–32, 149; legal theory of enterprise entity 138–41
risk: risk-bearing and management 217–19, 223, 228; risk-taking in classical economic theory 101–2; sharing of 35–36; unsystematic risk 333–35
Robbins, Lionel 86
Robé, J.-P. 359
Roberts, J. *et al* 356
Robinson-Patman Act 211
role playing 259
Romagnosi, Gian Domenico 112
Romano, R. 176–77n4, 360, 361
Romano, S. 296
Ross, S.A. 55, 322–23
Rouast, A. 140
Rowland, S.W. 86–87
Rowley, Scott 274, 276
Royce, Josiah 113
Ruben, David-Hillel 260n22, 277
Ruggles, Richard 74

Santa Clara v. Southern Pacific R.R. (1886) 161, 162
satisfactory return: idea of 202–3, 212, 213–14; profit maximization or 204–5
Sauvain, H. 320
Say, Jean Baptiste 107, 217, 220
Schär, Friederich 117
Schmalenbach, Eugen 5, 14, 111, 113–14, 120, 121, 126, 249; *Privatwirtschaftslehre* 114, 116–17
Schmidt, Fritz 118
Schmoller, Gustav von 127
Schumpeter, Joseph A. 15, 45, 109, 133–35, 137, 148–49, 217, 220, 224, 232n16
Schwayder, K. 117
scientific study, application of 120–21
Scott, A. Richard 243
Searle, John R. 281
Securities and Exchange Commission (SEC) 49n31, 172
self-interest of organisations 58
Senior, Nassau William 107
Shapley, L.S. and Shubik, M. 79
shareholder value *see* shareholders
shareholders: aspirations and corporate social responsibility (CSR) 154, 171, 174–75; entity-firm approach to firms 330; and firm, Miller-Modigliani proposition 323–24; primacy of 351–52, 353, 358; rejection of 361–62; responsibility of 96; shareholders equity interest and net income 256; shareholder-manager relations 23, 32–38, 35; shareholder value 353, 356, 370; shareholder wealth maximization (SWM) 321, 322, 324, 325, 329, 331–33, 337, 344n17; social-political factors and 101–3
Sharpe, W.F. 324, 333
Sherman Act 165–66, 211
Shiller, R.J. 327
Shleifer, A. and Summers, L.H. 327
Shubik, Martin 9, 11, 12, 13, 73–81, 109, 237, 238–39, 249, 257
Siegman v. Elec. Vehicle Co (1905) 165
Simon, Herbert A. 5, 6, 9–11, 13, 15, 17, 21–23, 27, 39, 54–72, 109, 237, 246–47, 281–82, 327; employment relationship proposed by 28–29; mythical Martian visitor idea 11, 56, 61, 68–69, 254–55, 257, 328
Sivan, Dganit 9, 14–15, 153–85, 272, 294, 308, xii
Smith, Adam 38, 107, 212
Smith, Thomas A. 272, 274
Smyth v. Ames (1898) 148, 151n20
social and economic relationships, interpretation of 3–4
social control, enterprise as object of 141–44
social institutions, assumption of fixity in 302–3
social organisation 32–33, 58
social responsibility 15; *see also* corporate social responsibility (CSR)
social sciences, methodologies of 112–13
Societas Europaea (SE) 368–69
Solomon, E. 331
Solomons, David 86, 87
Southern Railway Co v. Greene (1909) 164
special liability enterprises 192–93
Spencer, Herbert 119
stakeholders 35, 36, 47
statistics *versus* dynamics 329
Stauss, James H. 15–16, 216–34, 247, 273, 276, 278, 332
Steffens, L. 166
Stiglitz, J.E. 55
stock and securities markets 92, 101–3
stock-price mechanism 351–52
Stone, C.D. 366
Storey, R.K. 117
Stout, Lynn A. 273
Studenski, P. 73
sub-division of firms 186
subsidiary agreements 196–97
substance and form, concepts of 293
surplus, problems of 5
system, firm as 5–7, 9–11, 31, 39–43, 237–42, 246–48, 259, 266, 294
Swanson, Guy E. 281

Taylor, F.W. 111
Taylorism 44
team production 25, 36–37, 38–39
technostructure 350–51
Teubner, G. 366, 368
Thaler, R.H. 327
theory: 'adaptation' theory 24; *Betriebswirtschaftslehre* 114–16; economics and accounting, theoretical differences 79–80; of enterprise entity 186–200, 358; equilibrium theory 237–39; finance theory and the firm 318–23, 323–25; of firm 23–32, 82–91, 111–12, 128, 187, 240–42, 283–84; game theory 66, 79–80; general equilibrium theory 75; Hawkins-Simon

theorem 68; high theory in economics 4, 9, 10, 13–14; 'hold-up' theory 24; incomplete contracts 24–25, 26; integrated enterprise 305–6; of organisations 70; property rights, theory of 24–27, 29, 33–35, 37; 'rationing transactions' 135–36; real entity theory 175–76, 274–76, 285; revenue, theory of creation and distribution of 135–36; transaction costs theory (TCT) 24–26, 242–43; unitary business, theory of 305–6; valuation, theory of physical 148; *see also* agency theory; classical economic theory; firm, economic theory of; neo-classical finance theory (NFT)
Thirlby, G.F. and Buchanan, J.M. 86
Thomasson, Arnie L. 281
transaction costs 88–89; transaction costs theory (TCT) 24–26, 242–43
Truchy, Henry 140
Tuomela, Raimo 279
Thünen, J.H. von 107
Tyson, L.D. 172

*ultra vires* doctrine 164, 165, 166, 179n28
uncertainty: and continuity 337–38; coordination and 67; and entrepreneurial activity 230–31
undertakers 259
Union Pacific v. Rock Island R.R. (1896) 166
unitary enterprise: criteria for identification of 306; industrial groups as 304–6
United Kingdom *see* Britain
United States 3, 4, 14, 37, 49, 56, 86, 94–95, 100, 103
United States v. E.C. Knight Co (1895) 165–66
Unocal Corp v. Mesa Petroleum Co (1985) 170
Ure, Andrew 38

value-additivity 329
value based management 353, 354
value maximization, principle of 256
values: capital structure decisions and costs 336; financing values or costs 339–40; and NFT 326–27; revenues and costs 330–31; risk and equilibrium 333–35
Veblen, Thornstein 4, 111
Vernon, R. 172
Vining, Joseph 275

Walker, Francis A. 217, 220
Walras, Léon 75, 107, 108, 109, 135, 310
Walsh, J.P. and Avi-Yonah, R.S. 176
wealth holding 94–95
wealth maximization 331–33
Weinstein, O. and Azoulay, N. 42
Weinstein, Olivier 9, 10, 13, 21–53, 131, 283, xii
Weisbrod, B.A. 66
Weissman, David 276, 279
Wells, H.C.A. 154
Weston, F.J. 320, 331
White, J.B. 174
whole, firm as ; wholeness 5–6, 42, 120, 132, 242, 259, 270
Wicksell, Knut 127
Wieser, Friedrich von 107
Williams, C.A. 154, 176–77n4
Williamson, O.E. and Bercovitz, J. 360
Williamson, Oliver E. 5, 26, 27–28, 40, 55, 59, 126, 243, 268, 283, 298–99, 360, 361, 362
Windelband, Wilhelm 110, 113
Winter, S.G. 42–43
Wittgenstein, Ludwig 128
Wold, H.O. 77–78
Wood v. Dummer (1824) 177n14, 178–79n23

Zahn v. Trans-America Corp (1946) 194
Zappa, Gino 4, 5, 6, 14, 111, 114, 119, 120, 121, 123, 126–27, 249, 251, 298, 303, 308, 310
Zeckhauser, R. 326–27
Zingales, Luigi 29, 30, 36, 279, 322, 360, 361–62, 371n6