Economic and Social
Research Institute

EE

The Economics of an Ageing Population

MACROECONOMIC ISSUES

EDITED BY
Paolo Onofri

ESRI Studies Series on Ageing

The Economics of an Ageing Population

ESRI STUDIES SERIES ON AGEING

In April 2000 the Japanese government launched a series of comprehensive, interdisciplinary and international research projects called ‘the Millennium Projects’ and as part of this initiative the Economic and Social Research Institute (ESRI) of the Cabinet Office of Japan initiated a two year project entitled ‘A Study on Sustainable Economic and Social Structures in the 21st Century’, which focuses on ageing and environmental problems in the Japanese and international context.

The *ESRI Studies Series on Ageing* provides a forum for the publication of a limited number of books, which are the result of this international collaboration, on the three main issues of macroeconomics, pension and social security reform, and the labour market. The series is invaluable to students and scholars of public economics and public finance as well as policymakers and consultants.

Titles in the series include:

The Economics of Social Security in Japan
Edited by Toshiaki Tachibanaki

The Economic Impacts of Population Ageing in Japan
Landis MacKellar, Tatiana Ermolieva, David Horlacher and Leslie Mayhew

The Economics of an Ageing Population
Edited by Paolo Onofri

The Economics of an Ageing Population

Macroeconomic Issues

Edited by

Paolo Onofri

Professor of Economics and Director of the Department of Economics, University of Bologna, and General Secretary, Prometeia Associates, Italy

THE ESRI STUDIES SERIES ON AGEING

Edward Elgar

Cheltenham, UK • Northampton, MA, USA

Published by
Edward Elgar Publishing Limited
Glensanda House
Montpellier Parade
Cheltenham
Glos GL50 1UA
UK

Edward Elgar Publishing, Inc.
136 West Street
Suite 202
Northampton
Massachusetts 01060
USA

A catalogue record for this book
is available from the British Library

Library of Congress Cataloguing in Publication Data
The economics of an ageing population : macroeconomic issues / edited by
Paolo Onofri.
p. cm. — (The ESRI studies series on ageing)
Includes index.
1. Age distribution (Demography). 2. Economic development. 3. Pensions.
I. Title: Economics of an ageing population. II. Onofri, Paolo. III. Series.

HB1531.E27 2004
339′.084′6—dc22 2004042672

ISBN 1 84376 779 1

Printed and bound in Great Britain by MPG Books Ltd, Bodmin, Cornwall

Contents

Contributors

Masanao Aoki, University of California, Los Angeles, CA, USA.

Massimo Baldini, University of Modena-Reggio Emilia, and Prometeia Associates, Italy.

Paolo Bosi, University of Modena-Reggio Emilia, and Prometeia Associates, Italy.

Barry Bosworth, The Brookings Institution, Washington, DC, USA.

Ralph C. Bryant, The Brookings Institution, Washington, DC, USA.

Gary Burtless, The Brookings Institution, Washington, DC, USA.

Shin-ichi Fukuda, University of Tokyo, Japan.

Maria Cecilia Guerra, University of Modena-Reggio Emilia, and Prometeia Associates, Italy.

Fumio Hayashi, University of Tokyo, Japan.

Futoshi Kurokawa, University of Tokyo, Japan.

Carlo Mazzaferro, University of Bologna, and Prometeia Associates, Italy.

Warwick J. McKibbin, The Brookings Institution, Washington, DC, USA and Australian National University, Canberra, Australia.

Kazunori Minetaki, Fujitsu Research Institute, Tokyo, Japan.

Ryoko Morozumi, University of Tokyo, Japan.

Kiyohiko G. Nishimura, University of Tokyo, Japan.

Paolo Onofri, University of Bologna, and Prometeia Associates, Italy.

Edward C. Prescott, University of Minnesota, Minneapolis, MN, USA.

Claudia Sahm, University of Michigan, Ann Arbour, MI, USA.

Masato Shirai, Meiji Gakuin University, Tokyo, Japan.

Hiroshi Yoshikawa, University of Tokyo, Japan.

Preface

At the beginning of the twenty-first century the world must place the highest priority on constructing a sustainable socioeconomic system that can cope with the rapid ageing of populations in developed countries and with the limited environmental resources available in both developed and developing countries. At first glance, the problems of ageing and the environment may seem to be quite separate issues. However, they share a common feature: they both deal with intergenerational problems. The essence of the ageing problem is how to find effective ways for a smaller, working generation to support a larger, ageing generation. The crux of the environmental problem is to find a feasible way to leave environmental resources to future generations. Moreover, in terms of consumption, slower population growth may slow consumption and help to alleviate environmental problems. On the other hand, a rapidly ageing society may use more energy-intensive technology to compensate for the inevitable labor shortage, and thus cause deterioration on the natural environment.

Today, these concerns are highly applicable in Japan. The pressure created by the rapid ageing of the Japanese population is becoming acute; Japan must construct a sustainable society that does not create intergenerational inequity or erode the public welfare. At the same time, Japan cannot deplete its environmental resources and energy, as this would leave future generations with an unbearably heavy burden.

The Japanese government has recognized the vital importance of both problems. To explore and implement solutions for this difficult task, in April 2000 former Prime Minister Keizo Obuchi launched several comprehensive and interdisciplinary research projects known collectively as the 'Millennium Project'. As a consequence in the same month, the Economic and Social Research Institute (ESRI), Cabinet Office, Government of Japan, initiated a two-year project entitled 'A Study on Sustainable Economic and Social Structure in the Twenty-first Century'. While the Millennium Project covers a wide range of topics and disciplines such as natural science and technological innovation, the project conducted by ESRI places major emphasis on social science. While taking into account technological innovation and feasibility, it focuses on ageing and environmental problems. It aims to design a desirable socioeconomic structure under the pressure of an ageing population and environmental constraints by identifying the necessary policy tools to attain stable and sustainable growth.

This project is being implemented with close collaboration among Japanese as well as foreign scholars and research institutes. Besides Japanese scholars and institutes, foreign participants have been involved from, among others, the United States, the United Kingdom, Norway, Austria, Italy, Australia, Korea and Thailand. In total there are ten countries and 30 working groups.

AGEING SOCIETY

The Japanese population is graying rapidly, and the elderly are expected to account for a quarter of the nation's total population in 2020. Japan needs to reform its social security system, pension management schemes, financial/capital markets and labor markets if it hopes to create a better and sustainable future society. When looking at Japan's problems head on, we see giant fiscal deficits, bad-loan problems, and long-term structural problems that must be fixed aggressively or suffered indefinitely. The clock is ticking, and the time to act is short, but our problems are not insurmountable. Pessimism is not our rallying call. Rather, we are optimistic, and encouraged by the European example. In Europe, countries have worked hard to improve fiscal conditions and social security, and they have attained positive results and succeeded in uniting their separate markets.

Studies on ageing populations can be divided roughly into three categories: (i) macroeconomic problems related to the decline in the workforce due to ageing; (ii) social security systems, with many of the studies looking at pension systems; and (iii) the labor market, for example, employment of the elderly, competition with younger workers, the female workforce, and immigrant workers. We have made a specialized study of our theme from these perspectives.

Many people are pessimistic about the effect of ageing on the macroeconomy due to the reduction in the labor force. However, in this study, there were others who expressed a challenging view that various impacts, including those on economic growth, can be coped with fully and overcome by technological progress and other measures. Some strongly suggested that the current economic and social system would need to be reformed to achieve the flexibility required. Accelerated return to prosperity and the realization of economic growth at a comparatively high level would reduce costs and alleviate the distress that a change in the system in the transition period would involve, helping to carry forward the reform smoothly. In other words, measures to be taken to improve the present situation of the

Japanese economy do not differ significantly from what is required to build the future sustainable economy and society.

In order to emerge from the current stagnant state of the macroeconomy as soon as possible, it is necessary to reconstruct the financial system in such a way as would bring about efficiency of fund allocation and recover efficiency of the labor market. Some also pointed out that it will be difficult to achieve economic growth amid an increasingly ageing population unless a sound labor market and an efficient financial and capital market are established. Finding solutions for the bad loans and forming an efficient financial and capital market will not only make it possible to diversify portfolio selection and pension asset management of individuals but will allow enterprises to raise their funds efficiently. At the same time, there were many who expressed the opinion that the soundness of the Japanese economy would be recovered and a sustainable economy and society will be realized by defining public participation in the social security system (defining the roles which the public and the private sectors should play in social security) and building a safety net against various risks. As a timeworn story, it is pointed out that the formation of an efficient market through the improvement of various regulations and systems in the labor market will not only stimulate participation of the aged and women in the market and make up for the labor shortage but will also make a variety of employment forms feasible and contribute to the formation of a society in which people will feel that their life is worth living. There are also some who see the necessity of establishing the concept of equity in the social security system and employment of the aged since these factors bear a significant relationship to the age at which payment of pensions is started, sustainability of the system and fair sharing of the pension cost between the young and the aged.

RESOURCE AND ENVIRONMENTAL PROBLEMS

Studies on resource/environmental problems reflect a closed-loop model of the economy and society. These studies are divided into four themes: (i) studies on waste management, which cover a wide variety of empirical studies; (ii) studies on sustainability and technological innovation related to resources/energy; (iii) studies on potential policies for addressing changes in climate; and (iv) studies on the relationship between environmental policies and economic policies, including employment policies.

From the standpoint of long-term sustainability of global resources, official involvement including policy measures and the development of new technology to remove environmental restriction will be called for. There is a high possibility that new technology creates new products, stimulating demand, developing new industries with high productivity, and bringing about a renewed sustainable economic growth. New technology in the 21st century should contribute to the construction of the closed-loop economy and society and enhance resources and energy efficiency, properly dispose of waste and increase efficiency of reusing resources. Furthermore, it must generate renewable energy efficiently and on a large scale. It may sound paradoxical, but past experience suggests that various restrictions imposed on economic activities, and the existence of regulation actually stimulates new technological development which helps to break through the restrictions.

Today, environmental issues encompass a very wide range of problems from territorial disposal of waste to the global environmental issues. It has been pointed out that policy mix ingeniously combining such methods as regulatory and economic instruments, voluntary agreements and international emissions trading is essential for coping with these issues. What is important then is a policy which skillfully uses incentives, making use of market mechanisms. Japan is one of the most advanced countries in environment-related technology in the world. However, further technological breakthroughs will be called for in the future and it has been pointed out that use of market incentive and official support is essential in the fields where long-term risks are uncertain.

It is suggested that concrete behavior by a community based on shared information (bottom-up approach) and a change in behavior of individual as consumers provide one of the keys in the closed-loop economy and society. Such change in the behavior of individuals is caused by the diffusion and permeation of concepts such as precautionary principle, one of the environmental principles in the EU.

Addressing the construction of the closed-loop materials-cycle economy and society is an effort that is particularly called for in the Japanese economy which is faced with restrictions on resources and energy and it may be said to be a way by which we must seek sustainable progress and growth of a Japanese type. Furthermore, some pointed out that, for its climatic and topographical conditions and for its population density, the waste management system in larger cities of Japan can be a model for Asia and Japan will play an important role in the Asian area in coping with global environment issues.

SYNTHETIC CONSIDERATION

None of the various issues dealt with in this project are independent issues. They are issues closely related to each other and they will require simultaneous decision-making, all of them when a desirable scenario is to be visualized. And both issues of ageing and environment will have a strong impact not only on the welfare of our generation but also on that of future generations. An awareness of issues that are common to the ageing population and the issue of environment is the context in which we should make use of market mechanisms.

There is a case for saying that we should make good use of market mechanisms by utilizing economic means in order to efficiently attain goals in health care and annuity, employment of the aged, women in the labor force, climatic change and waste disposal. However, many of these issues are examples of market mechanisms not functioning efficiently, or what one calls a failure of the market. Utilizing market mechanisms in this domain involves various difficulties such as internalizing externalities and we cannot avoid classical problems such as efficiency and equity and what roles should be played by the public and private sector.

Research into the ageing population and the environment has thrown into relief the importance of technological innovation. Looking back on human history, we find that restrictive conditions gave birth to technological innovation which presented a way out of the difficulties faced. When we think the 21st century has severe restrictions imposed upon it, such as resources, energy, labor force and population, we may take a positive view of things and say that, reversely, the century carries full conditions that will give birth to technological innovation. If we could succeed in achieving technological innovation in this global situation, formation of more enriched economy and society would be materialized.

In the short term, countries with ageing populations will likely elect to consume more energy – without any thought to the long-term impact of their consumption patterns and economic activity. This tendency must be offset with a new sense of sustainability, one that looks to the future, one that thrives on improved resource/energy efficiencies, one based on eco-friendly waste disposal and new eco-friendly technologies. Without lowering living standards, we must solve global environmental problems, and overcome the constraints of limited energy resources. To do so will require the creation of a closed-looped economy/society. Failure to do so may spell the end to our way of life in the not-so-distant future.

In this project, we explore optimal solutions to social optimization problems. After taking into account the political and social constraints we face, and after alignment and coordination with the results of the studies, we will sketch out an ideal design and examine the possible direction of future research.

This project came to an end in March 2002. It solved many theoretical and empirical issues, but has created new debates. Twice a year, all the members of the project, along with selected participants, met to discuss the results of the research. Regretfully, it has not been possible to reproduce the fruitful discussion in this volume.

Overall, the papers presented in the project were extremely challenging, and covered a wide range of topics. In the near future we strongly hope we will have a chance to discuss the research once more from a common standpoint.

The result of this research appears in print from Edward Elgar Publishing Ltd as part of the ESRI study series, available to policy makers, academics and business people with a keen interest in these subjects. The series on ageing problems covers macroeconomics, social security and the labor market. Unfortunately, because of space limitations we regret that we are able to publish only selected papers from the total research effort. The research papers to be published were selected by the Editorial Board members. We would like to acknowledge the ceaseless efforts of the members of ESRI throughout the project period, especially those of the Department of Administration Affairs. Last but not least, we would like to thank Dymphna Evans, Matthew Pitman and Karen McCarthy from Edward Elgar Publishing.

Yutaka Kosai, President, ESRI

Acknowledgments

Chapter 2 is reprinted from *The Review of Economic Dynamics*, **5**(1), January 2002, pp. 206–35, Hayashi and Prescott: 'The 1990s in Japan: A Lost Decade', with kind permission from Academic Press, an imprint of Elsevier Science.

Large parts of Chapter 3 are reprinted from *Asian Economic Papers*, **2**, 2003, pp. 85–136, Nishimura and Shirai: 'Can Information and Communication Technology Solve Japan's Productivity Slowdown Problem?', with kind permission from MIT Press.

Chapter 4 is reprinted from the *Journal of Economic Behaviour and Organization*, **48**, 2002, pp. 127–54, Aoki and Yoshikawa: 'Demand Saturation–creation and Economic Growth', with kind permission from Elsevier.

Introduction

The fear that market economies would stagnate had been present in economic thought for almost a century when, after the First World War, the decline in the rate of growth of the populations of the then more advanced economies brought it back to centrestage. In the 1920s, although migration had fallen to between a third and a quarter of the flows of the pre-war decades, the rate of growth of the population of European countries almost halved. The acceleration[1] of the long-run decline in the birth rates was the main factor. The parallel decline in the death rates was not enough to compensate for its effects. The coincidence, in the next decade, of the lower population growth with the Great Depression[2] and with the development of the analysis of the non-competitive market structures, justified the new stagnationist approach to long-run growth. This revival prolonged itself into the economic literature of the 1950s.

In the meantime, two demographic shocks of different sign and size hit the more advanced economies: some of them experienced a further sharp drop in the birth rate during the Second World War, for obvious reasons; all of them reacted to the end of the war with a huge and long-lasting acceleration in the birth rates. This latter reaction, of the opposite sign to that observed after the First World War, was accompanied by the so-called 'golden age of growth' of the industrialised world, and a stagnationist approach was no longer mentioned.

Viewed under this perspective, the baby boom at the end of the 1940s and in the 1950s may look like an over-reaction along the long-run declining trends of both birth and mortality. Such a long over-reaction suggested that there are more interactions among economic growth, social policies and population dynamics than previously thought.

Nowadays, we might also wonder whether the two or three decades of rapid growth after the Second World War were an outlier in the very long-run process of growth; whether the current deceleration of the population growth in Japan and in Europe is the reason for their lower rate of growth; whether, in more general terms, the expected decline in population in future decades will necessarily imply a sharp reduction in the rate of growth of the industrialised economies.

The pendulum of the attention of the economists has moved back to the relationship between demography and growth. Numerous contributions have been published on the question whether the phase of *demographic transition* to a less numerous and older population implies systematically lower rates of growth. This book, on population ageing and growth, tackles some aspects of this problem from both a theoretical and an empirical point of view as part of a larger project aimed at explicating the possible effects of an ageing population on the macroeconomic performances of social security, the labour market, and growth. Two other books in the series have focused on population ageing and social security, and on population ageing and the labour market.

The macroeconomic questions raised by the ageing of a population are well known and shared; the answers are much less so. This book adds new contributions to the spectrum of answers, based mainly on empirical findings deeply rooted in theoretical analysis. The main focus is on Japan, but we also turn our attention to Italy and the United States.

AGEING AND THE MACROECONOMIC ISSUES

Labour Supply

The decline in population implies a reduction in the supply of labour, and a change in its composition. The reduction in the actual supply of labour is expected to be offset by an increase in the participation rates of the working-age population. This is strictly tied to the likelihood of an increase in female labour force participation, and the participation of both male and female workers within the 55–65 age bracket. Wage incentives and appropriate social policies are required. Given that younger female cohorts are spontaneously participating to an increasing extent, a more rapid increase in the average participation rate would depend mainly on policies aimed at changing the attitude to work of the more mature female labour force. These changes face cultural obstacles and mismatches in the demand and supply of skills, which cannot be ignored. If these policies are not successful the only tool to increase labour supply is to allow higher rates of immigration.

The change in the composition of the labour supply is a question relating to the impact of a relatively older workforce on average productivity. The age–earnings profile that we have inherited from the past shows an increase in earnings as workers age, suggesting that productivity increases as individual workers age. Does the increasingly larger share of mature workers that we expect for the future really imply a higher average output per employee? If the answer is in the affirmative, it would compensate, at least partially, for the reduction in absolute value of the potential labour force; but this is open to question. For example, consider the features of the new wave of technical progress through information and communication technology (ICT). These require strong efforts in the lifelong learning approach to human capital maintenance in order to allow workers to be able to fully use the new technologies. Are the industrialised countries ready to implement such human capital policies? Is the efficiency of the new production processes still based mainly on the experience the workers accumulate throughout their working life, as in the more traditional manufacturing activities? The possible mismatch between the competence required by the application of ICT and the ability of an older population to keep up with those technologies, will influence industrial relations, wage policies and, more generally, the dynamics of total factor productivity (TFP).

There are not only supply-side effects of a declining population. We should not ignore the possibility that there may also be demand-side effects, such as reduced stimuli to radical product innovations, which would reduce the rate of invention of new products, and the frequency of starts of new product cycles. This reduced rate of product innovation might not compensate for the saturation of the more mature product cycles. If we view the long-run growth process as the result of subsequent starts of different product cycles, the reduction in the start rates implies a reduction in the long-run growth.

Capital Accumulation and Technical Progress

Let us consider the supply effects. If the higher intensity of utilisation of the working-age population, and the possible better quality of the labour supply is not enough to compensate for the decline in the level of potential employment, the expectation that technical progress will expand the production possibility frontier that demography narrows would rely only on the quantitative and qualitative aspects of the capital accumulation process, and on the nature of technical progress. If more capital-intensive production processes are required to help productivity growth to compensate for the decline in the labour input growth, the fact that the reduction in the size of the labour force releases resources should not be ignored. Evaluating the actual amount of resources released requires considering the behaviour of the capital depreciation rate: the more rapid the technical progress, the shorter the average economic life of capital and the fewer the resources released by the decline in employment.

The other side of the coin is the spontaneous effect of ageing on saving propensity. The decline in the fertility and mortality rates have opposite effects on the dependency ratio. Currently, the dependency ratios in the industrialised countries are more or less the same as they were a century ago. Of course, the crucial difference is the relative size of young and old-age dependency. In future decades, the strong increase in the elderly dependency ratio will substantially outweigh the reduction in the young dependency ratio, even though two factors are counteracting the downward trend in the young dependency ratio. First, at the current immigration rates, the reduction in the fertility rate seems to have reached its floor, at least in the fastest-ageing countries such as Japan and Italy.[3] Second, the young are systematically delaying the age of entry into the labour market.

The young and elderly components of the dependency ratio are expected to have different effects on saving rates through the life-cycle planning of consumption and the attitudes towards the future expressed by the intertemporal preferences of the economy as a whole. Lower young dependency is very likely to imply a reduction in the propensity to consume. Higher elderly dependency implies either a higher propensity to consume, if we consider only the effect of asset decumulation, or a lower propensity to consume if we consider the strong precautionary motivation for saving that elderly people tend to show, in view of their longer life expectancy as retirees, and the implicit higher probability of old-age disability. Moreover, it is likely that a more mature working-age population, expecting to live longer, will save at a higher rate. As a whole, population ageing is likely to exert a larger number of effects with a positive sign on saving, but we know very little about the relative sizes of the negative and positive effects.

If the empirical answer is that the positive spontaneous impacts of ageing on saving are not enough to compensate for the negative ones, should governments enact policies aimed at stimulating saving or should they act to stimulate productivity growth directly? The question is even more intriguing if we consider technical progress to be endogenously determined. This alternative nature of technical progress allows different intergeneration distributions of the burden of capital accumulation. In other words, should the current generations bear the burden of reducing the social and political impact of the future intrageneration redistribution of consumption implied by the ageing of the population, or could it be spread more evenly over current and future generations?

Pension System Reforms

In the absence of a clear answer to this question, governments, with different degrees of awareness, seem to behave according to the following reasoning: in the next decades, population ageing will bring about, *ceteris paribus*, a reduction in consumption per capita with respect to output per employee; as a consequence, the ultimate goal of the pension reforms is to reduce the ratio of the average benefit to the average output per employee; this aim can be pursued through a reduction in the entitlements, changing the rules of the benefit determination, and/or an increase in the 'size of the cake', increasing the relative weight of the funded pillar of the pension systems *vis-à-vis* the unfunded one; this change is expected to increase the propensity to save, hence the long-run productivity growth.

If the increases in the propensity to save and the output per employee are not significant, the redistributive tensions will show up along two different channels. First, countries with a large share of the unfunded pillar will face political and social tensions; they will need a parliamentary vote in order to redistribute consumption, but parliaments will by then comprise a larger share of representatives of the elderly part of the population. Second, countries giving importance to the funded pillar will face a downward shift in the price of assets held by the elderly, in order to adjust the value of the asset decumulation required by an ever-increasing elderly population to the value of the new saving of a dwindling younger generation.

In other words, if the policies aimed at increasing both labour force participation and productivity growth fail, then the demographic shock translates into the choice between a political risk and a financial market risk. Are they diversifiable risks, or are they two corner solutions? The current common opinion seems to be based on the second assumption. At least in principle, financial markets allow returns higher than the rate of growth of the economy, so we should put all our eggs in one basket; this, moreover, would allow the diversification of the demographic risk through the financial and direct investment in the emerging countries still lagging behind in their demographic transition.

This is a very abstract solution. Countries with a largely pay-as-you-go system are already far ahead in the process of ageing and the cost of withdrawing from such a system would be very high. Thus a diversification of the two risks is more likely. On the one hand, governments will have to accept an increasing share of immigrants who will be fully integrated into domestic institutions. On the other, the residual impact of the consumption redistribution will be spread throughout the international arena through deficits in the foreign current account, compensated by outflows of capital that react to the domestic returns reduction. This international redistribution of production activity will be reflected in a GDP growth that is smaller than than GNP growth.

THE MAIN FINDINGS OF THE BOOK

Nothing particularly new has been added to the background situation as outlined above. What is really new is the effort made by the contributors to this book to draw answers to the questions raised from a well-detailed empirical analysis. Part I is devoted to the analysis of the long phase of stagnation of the Japanese economy. In Chapter 1, 'Economic growth under the demographic transition: a theory and some international evidence', Fukuda and Morozumi move from the working assumption that the rapid growth of the East Asian countries in the past decades has been supported by the sharp decline in the birth rate, which has reduced the dependency rate, hence increased the propensity to save. Now the the birth rate has stopped declining, whereas mortality rates are expected to continue decreasing, thus increasing the dependency ratio. The authors build an overlapping-generations model, where the young workers' propensity to save depends on their life expectancy. The theoretical result is that the strength of the statement that demographic decline implies growth rate decline is reduced if the young save more and feed a more intensive capital accumulation. This result granted, the authors challenge their theoretical conclusion on empirical grounds. Econometric cross-country estimates do not contradict their working assumption: the rate of growth of the population has a negative impact on the rate of growth of the economy, but the old-age dependency ratio has a positive effect, and the young dependency ratio a negative one. Moreover, the cross-country behaviour seems to suggest that if private saving is invested either abroad or in the financing of the larger public budget deficit, the positive impact on growth is reduced. As a consequence, the prospective impact of demographic transition on the economic growth of the Japanese economy will be the result of the balance between those two opposing effects.

In Chapter 2, 'The 1990s in Japan: a lost decade', Hayashi and Prescott take a different point of view. Two main factors have been at work on the supply side of the Japanese economy during the 1990s: the decrease in the rate of growth of TFP, and the progressive reduction in the working time of Japanese workers (from 44 hours a week to 40). The authors check their working assumptions through the calibration of a neoclassical growth model on the data of the pre-1990 Japanese economy, and show that their model predicts the slowdown of growth that has been experienced in the 1990s, the increase in the capital output ratio, and the decline in the rate of return to capital. The relevance that their model gives to TFP in explaining Japan's growth decline suggests that, in order to support growth, Japan does not need to increase the capital to labour ratio, hence does not need to increase the saving rate. However, it does need to improve the efficiency of the overall economic system, by reducing the area of rent positions sheltered from market competition, and the amount of subsidies paid to inefficient firms and to declining sectors.

Whereas Hayashi and Prescott identify the diminishing contribution of

TFP as a factor in the growth of the Japanese economy, in Chapter 3, Kurokawa, Minetaki, Nishimura and Shirai ('Effects of information technology and ageing work force on labour demand and technological progress in Japanese industries: 1980–1998') try to show why TFP growth has diminished. First of all, they build an extensive database for both the ICT capital stock of the different industrial sectors, and the labour inputs, disaggregated with respect to age and education. On the basis of their extensive empirical analysis they conclude that the new ICT capital stock services are substitutes for young workers with low education, and complements of workers with high education. Older workers with low education have become to a larger extent a quasi-fixed factor of production. With the new ICT technology, the long on-the-job experience of the older workers no longer represents a comparative advantage for the quality of Japanese products. The empirical analysis concludes that the productivity slowdown is the result of the inflexibility of older workers, the obsolescence of the management *vis-à-vis* the technology evolution, and the small externality effects of the new ICT that these factors imply. In other words, the change in the composition of the labour force, implied by the ageing of the Japanese population, does not fit too well the observed change in technology.

In Chapter 4 ('Demand saturation – creation and economic growth'), Aoki and Yoshikawa look for an explanation of the slowdown of growth in Japan not on the supply side, but on the demand side. Technical progress behaviour should not be approached through the concept of TFP. Technical progress proceeds in steps marked by innovations capable of creating demand, which validates the introduction of new products. As the new products mature, demand for them tends to reach saturation point, and the return to capital invested in those industries diminishes even if the quality of the mature products is improved and their production processes restructured. The rapid growth in the East Asian countries in the 1990s, and in Japan in the 1950s and 1960s, was a result of the appropriate combination of income distribution and new product introduction. In a sense, the authors try to map the Keynesian principle of effective demand onto the domain of medium–long-run economic growth. Whether the ageing taking place in Japanese economic society has something to do with the saturation of demand for current products or the reduction in new ones is a question which is alluded to, but not pursued.

Part II comprises three chapters dealing with the impact of demographic transition on the different kinds of pension systems, and the possible effects of their reform for income distribution and growth. In Chapter 5, ('Distributional impact of social security reform') Bosworth, Burtless and Sahm use a small neoclassical growth model combined with a microeconomic model based on several different age profiles of individual earnings to simulate both the distributional and scale effects of different kinds of reform of the pension system. Those effects are simulated in two different contexts (static and dynamic), and are evaluated in terms of lifetime income. The static simulations merely show the distributional impacts of the reforms among both young and old generations, and workers with high or low age profiles of earnings. The dynamic ones allow the induced higher savings and the implied increase in the resources available for future generations to be measured. Higher payroll taxes on the current generation, lower future benefits, and a larger share of total retirees' income due to the return on individual retirement accounts are the candidate reforms capable of increasing national savings through higher private or government savings. Higher national saving implies a more rapid growth of real wages (through an increase in the capital to labour ratio), but a reduction in the rate of return to capital, which changes the proportion of total retirees' income coming from either social security or individual accounts, given the different distributional rules: defined benefit versus defined contribution.

The different performances of the funded and unfunded pension systems depend on the different return to social security and to financial assets. They are both subject to risk: first a social and political risk, more uniformly spread over a relatively higher number of generations; the second, a market risk, possibly more concentrated on particular generations. In Chapter 6 ('Asset accumulation and retirement income under individual retirement accounts: evidence from five countries'), Burtless considers the risk component associated with the defined-contribution individual retirement accounts; the previous chapter ignored this factor. The main financial risks refer to the behaviour of the rate of return during the period of asset accumulation, to the rate at which the accumulated assets are transformed in pension annuities, and to the inflation risk during the retirement years. Burtless shows that the return on individual accounts can be a good one, but with a large variance of possible outcomes, according to the cohort the worker belongs to. The wide spread of results is based on the computation of the return the workers would have, had they invested for 40 years in individual accounts starting at different dates between 1927 and 1962. These computations apply to different financial portfolios, and to different industrial countries. To give a very simple example of the risk associated with the individual accounts, a Japanese worker who has invested in a conservative portfolio (50 per cent equities, 50 per cent bonds) for 40 years, and retires in a given year in the 1990–2002 period, may have a replacement rate of pension to his/her peak wage of between 24 and 103 per cent.

However, economic policy for ageing societies aims at a more differentiated pension system, reducing the promised benefits, increasing contributions and offering options for withdrawing from pay as you go. The hope is for a positive effect on saving, which supports higher growth. The Italian experience is quite interesting from this point of view. Baldini, Bosi, Guerra, Mazzaferro and Onofri deal with the process of reforms of the public pension system enacted in Italy between 1992 and 1997, in Chapter 7 ('Pension reforms, tax incentives, and saving in Italy'). They analyse the effects of those reforms on the individual behaviour of households, and conclude that, apart from a stronger short-run impact, the compensating increase in private wealth was about half the size of the change in net social security wealth induced by the reforms. This compensating increase is, however, concentrated among the middle and older age groups, while the young do not seem to have reacted significantly to the reforms, even if they are the generations most affected. Actually, the replacement rate in Italy is still relatively high, and the workers benefit from a system of compulsory precautionary saving. Contrary to the idea that pension funds do not take off owing to a lack of tax incentives, the authors show that the incentives are strong enough, even in an international comparison with the United Kingdom or the United States.

Independently of the financial risk associated with the asset value of the funded pillar of the pension system, ageing is expected to imply a reduction in the return to domestic capital investment; this possible general drift is likely to be contrasted by international capital movements towards a higher return in the 'younger' countries. Part III is devoted to the analysis of the influence of demographic transition on international trade, capital markets, and exchange rates. In Chapter 8 ('Incorporating demographic change in multi-country macroeconomic models: some preliminary results') Bryant and McKibbin insert demographic components (birth rates, mortality rates, age–earnings profiles) into a macroeconomic model of two open economies: a simplified version of the world economy. The agents of each region of this world economy model behave according to the life-cycle theory of consumption and capital accumulation, where demographic variables influence human capital. The authors simulate the effects of symmetric or asymmetric demographic shocks. An asymmetric demographic shock like the post-Second World War baby boom years, followed by the progressive ageing of the population, produces a long phase of higher growth (hence higher consumption per capita), higher interest rates and depreciation of the currency of the country hit by the shock. That phase is followed by another long period of adjustment to the baseline path, where growth (and consumption per capita) slows down, interest rates decrease and the currency appreciates. The shock takes about a century to be reabsorbed. The path of absorption of the shock is quite different according to the role that the age–earnings profile is allowed to play in determining the human capital evolution.

SUMMARY

The positive shock on the birth rate in the decades after the Second World War went hand in hand with the most important phase of growth ever experienced by the industrialised countries. Japan and Europe are currently experiencing the incipient phase of demographic transition after the baby-boom shock, associated with a decade of stagnation in Japan and a slower medium-term growth in Europe. In both cases, the demographic explanation is likely to be deeply rooted in the background, but it does not seem to be the dominant one. As for the future, given the expected size of the ageing process, demographic transition is likely to work as a more important attractor for the economic growth of the most rapidly ageing countries. The chapters presented in this book show that this does not necessarily imply a long-run wave of dramatic decline in growth for those countries. Immigration and integration of foreign workers, deeper integration of capital markets with agreed mechanisms to reduce their volatility, higher domestic participation rates and, more relevant, a stronger rate of increase in technical progress would allow the strength of the demographic attractor to be reduced.

The chapters presented in this book suggest that economic policy has to face two different kinds of challenge. First, how to let the various factors work out all their positive effects. As for the increase in the participation rate, there is general agreement both on the goal and on the measures to be taken. It is likely to be a medium-run process, emerging as the younger cohorts age. There is less agreement on how to support higher long-run growth. More market efficiency is necessary to allow productivity to accelerate, but is it enough? An ageing population is bound to imply an increase in the propensity to consume, even if, as the contributions to this book argue, to a lesser extent than is usually thought. Nobody denies that long-run economic policy should aim to reverse the decline in the propensity to save; nonetheless, in countries like Japan, Italy and Europe in general, where populations are already ageing, there are strong doubts about the opportunity to implement such policies in the short and medium terms. The second kind of challenge is how to make the redistribution of consumption, which the ageing process necessarily requires, socially and politically viable. There is general agreement on the transformations required by the current pay-as-you-go pension systems to make them more equitable. Less agreement can be found about the extent to which they are to be transformed into individual funded accounts. The contributions to this book supply tools for the analysis of the redistributive impact among and within the generations. However, they cannot say anything about whether those measures will actually keep potential social and political tensions under control.

NOTES

1. See Angus Maddison, *The World Economy: A Millennial Perspective*, Paris: OECD, 2001.
2. See Joseph A. Schumpeter, *History of Economic Analysis*, New York: Oxford University Press, 1954, vol. III, ch. 6, § 1.b.
3. This does not prevent the birth rates from further declining during the transition to a lower share of 25–45 year olds.

PART I

Demographic Transition and the Impact on Growth: The Case of Japan

1. Economic growth under the demographic transition: a theory and some international evidence*

Shin-ichi Fukuda and Ryoko Morozumi

1. INTRODUCTION

Recent studies based on cross-country regressions establish the existence of strong linkages between demographic change and economic growth.[1] In particular, studies such as Higgins and Williamson (1996, 1997), Bloom and Williamson (1998) and Bloom et al. (1999) show that demographic variables have played a large role in East Asia's economic success. According to these studies, the demographic transition – a change from high to low rates of mortality and fertility – has been more dramatic in East Asia during the twentieth century than in any other regions or historical period. A rapid decline in fertility, induced partly by the region's economic success, led to a substantial reduction in the youth dependency ratio, thereby helping to boost saving rates and rates of economic growth in the region. East Asia thus has had exceptionally favorable demographic characteristics in the form of high life expectancy and low fertility, despite its initially low income level. These studies, however, point out that the favorable demographic characteristics have a purely transitional effect on economic growth; this effect operates only when the dependent and working-age populations are growing at different rates. Therefore, they predict that economic growth in East Asia will likely slow in the future, because of stabilization of fertility rates at their current low levels and increases in the dependency ratio as the population ages.[2]

In terms of demographic changes, Japan is a leading East Asian country that has enjoyed the transitionally favorable demographic characteristics earlier than other East Asian countries. During the past 40 years, Japan has enjoyed a gift from a demographic phase in which the youth dependency burden decreased and the proportion of the working-age population increased. The result was high saving rates and sustained economic growth abetted by demographic forces.[3] Population is, however, expected to peak in 2007 and the share of those aged 65 and older in the total population will rise about 10 per cent in the next quarter-century. The labor force, on the other hand, will decline by 0.6 per cent per year. The above prediction thus implies that the declining labor force will bring about slower economic growth in the near future.[4]

However, how serious the problem will be crucially depends on what determines the potential growth rate of the economy. From the point of view of traditional 'growth accounting', the prediction that the declining labor force will lead the economy to slower growth is actually not self-evident. Most of the results in growth accounting have found that capital accumulation and total factor productivity (TFP) growth are much more important factors than labor force to explain the growth rate of the economy. The finding is qualitatively robust in many developed countries. We must inquire, therefore, how large capital accumulation and TFP growth will be in the near future as the population ages in Japan.

In order to answer this inquiry, this chapter first considers an overlapping-generations model with capital accumulation and uncertain lifetime horizon. In the model, the saving rate of the young is larger when life expectancy is larger or when the rate of population growth is smaller. The model thus suggests that a rise in the share of the retirement-age population might enhance economic growth through increasing capital accumulation per worker.

In the second half of the chapter, we examine this hypothesis by cross-country regressions. The estimations of cross-country growth equations are standard in the literature (for example, Barro 1991; Levine and Renelt 1992; and Mankiw et al. 1992). Most of the previous studies derived the equations from neoclassical growth models (for example, Solow 1956). In contrast, we derive the corresponding equation from an overlapping-generations model. The derived equation satisfies the standard convergence property. It also formulates the demographic impacts on economic growth based on our model specification. In the estimations, several demographic variables are highly significant. In particular, we find that given other variables, the share of the retirement-age population has a significantly positive impact on per capita GDP growth. The result suggests that a decline in the working-age population does not necessarily slow down the rate of economic growth.

A crucial point in our analysis is that higher life expectancy and lower population growth increases the saving rate of the young. The negative impact of a decline in the working-age population may thus be dominated by the positive impact through capital accumulation. The implication is never derived in the Solow-type growth models where the saving rates are exogenous. Our theoretical implication is, however, consistent with models of Ehrlich and Lui (1991) and Kalemli-Ozcan et al. (2000) who showed that population ageing has a positive impact on capital accumulation.[5] These studies focused on the role of human capital accumulation rather than physical capital accumulation as an engine of economic growth. More importantly, these studies did not present empirical evidence to support their theoretical results. Allowing uncertain lifetime horizon, we formulate an overlapping-generations model in an empirically tractable way and test its relevance by cross-country regressions.

Almost 40 years ago, Coale and Hoover (1958) proposed their famous dependency hypothesis. According to this hypothesis, some of the impressive rise in Asian saving rates can be explained by the equally impressive decline in dependency burdens and some of the booming East Asia phenomenon can be explained by the resulting savings boom. The hypothesis, however, implies that the demographic transition is eventually manifested in a big elderly burden, resulting in a deflation of the miracle. It thus predicts that sustained economic growth in East Asia should tend to vanish as the elderly dependency rate rises over the next three decades. Our empirical results, however, suggest that the prediction is not necessarily correct in East Asia. Through increasing the saving rates of the working-age population, a rise in the elderly dependency rate may enhance economic growth if the savings are efficiently invested for domestic capital accumulation.

The chapter is organized as follows. Section 2 explains the analytical framework and Section 3 discusses the impacts of demographic factors on economic growth in our theoretical model. Section 4 specifies a basic equation for cross-country regressions. Section 5 presents the main empirical findings. Section 6 presents alternative estimation results and Section 7 compares our results with previous studies. Section 8 extends the analysis to the case where international capital flows exist. Section 9 discusses implications for East Asia and Japan. Section 10 summarizes our main results.

2. THE ANALYTICAL FRAMEWORK

In this section, we present a simple overlapping-generations model where demographic factors have significant impacts on economic growth. The analytical framework is an extension of Diamond (1965) to the case where the lifetime horizon is uncertain. The economy is composed of consumers and firms. In each period, a new cohort of consumers is born. Each consumer lives for two periods with probability q but dies at the end of the first period of life with probability $1-q$. The number of consumers born at time t is N_t. Population of the young generation grows at rate $n \equiv N_t/N_{t-1} - 1$. The ratio of the retirement-age population to the working-age population,

$g \equiv qN_{t-1}/N_t$, is thus equal to $q/(1+n)$ for all t. Since population in period t is $N_t + qN_{t-1}$, the rate of population growth in period t is equal to n for all t.

Consumers work only in the first period of life, supplying inelastically one unit of labor and earning a real wage of w_t. They consume part of their first-period income and save the rest to finance their second-period retirement consumption. Since there is individual but no aggregate uncertainty, there is scope for insurance in the model. Consumers can contract to have their financial assets v_t go to the insurance company contingent on their deaths. Assuming free entry and zero profit in the insurance industry, the insurance company will pay them a premium of $[(1-q)/q]v_t$ if they are alive in the second period of life. The insurance market is, however, highly limited in most of countries, particularly developing countries. In the following analysis, we therefore assume that only a fraction $1-f$ of their financial assets is insured. In the absence of insurance, consumers would die leaving unintended bequests with probability $1-q$. The amount of unintended bequests per worker at the beginning of period t is equal to

$$b_t = (1-q)\, r_t f\, s_{t-1}/(1+n), \tag{1.1}$$

where s_{t-1} is saving per worker held from period $t-1$ to period t and r_t is its (gross) interest rate.

A consumer born at time t consumes c_{1t} in period t and if any, c_{2t+1} in period $t+1$. At the beginning of period t, each consumer maximizes the following expected lifetime utility:

$$ln(c_{1t}) + q\beta\, ln(c_{2t+1}), \tag{1.2}$$

subject to the budget constraints

$$c_{1t} + s_t = w_t + b_t \text{ and } c_{2t+1} = r_{t+1}[1 + (1-f)(1-q)/q]s_t, \tag{1.3}$$

where β is a discount factor.

The first-order condition leads to:

$$s_t = \frac{q\beta}{1+q\beta}(w_t + b_t). \tag{1.4}$$

Equation (1.4) implies that the saving of the young is proportional to $w_t + b_t$ (his or her wage income plus unintended bequest) and is independent of the interest rate r_{t+1}. The result arises because the consumer's utility function is logarithmic.

Firms act competitively in the product and factor markets. Let Y_t be output at time t, K_t be physical capital stock at time t, and N_t be labor input in period t. We assume that firms have the production function $Y_t = AK_t^{\alpha} N_t^{1-\alpha}$, where $0 < \alpha < 1$. The term A is the level of exogenous technology. Define per worker output by $y_t \equiv Y_t/N_t$ and the capital–labor ratio by $k_t \equiv K_t/N_t$. The production function is then rewritten as

$$y_t = A\, k_t^{\alpha}. \tag{1.5}$$

Assuming that equation (1.5) denotes the production function net of capital depreciation, we can derive the equilibrium wage and (gross) interest rate in period t as follows:

$$w_t = (1-\alpha)\, A\, k_t^{\alpha} \text{ and } r_t = \alpha\, A\, k_t^{\alpha-1}. \tag{1.6}$$

The saving of the young in period t generates the capital stock that is used to produce output in period $t+1$. Since n is the growth rate of working-age population, this implies that $(1+n)\, k_{t+1} = s_t$. Because (1.1) implies that $b_t = (1-q) f r_t k_t$, it thus holds that:

$$k_{t+1} = \frac{q\beta}{1+q\beta} \frac{1-\alpha(1-f)-qf\alpha}{1+n} A\, k_t^{\alpha}. \tag{1.7}$$

Equation (1.7) describes the dynamic behavior of per worker capital stock in our model. It satisfies the stability condition for any positive initial capital stock.

3. THE IMPACTS OF DEMOGRAPHIC FACTORS ON ECONOMIC GROWTH

Impact on the Steady State of Per Worker Capital Stock

In this section, we investigate what linkages demographic factors have with economic growth in our overlapping-generations model. We first focus on the linkages of demographic factors with the steady state of per worker capital stock. Since the steady state corresponds to $k_t = k_{t+1}$, the model has a unique steady equilibrium such that:

$$k^* \equiv \left[\frac{q\beta}{1+q\beta} \frac{1-\alpha(1-f)-qf\alpha}{1+n} A \right]^{1/(1-\alpha)}. \tag{1.8}$$

Equation (1.8) implies that two demographic factors, q and n, have significant impacts on the steady-state value of per worker capital stock k^*. The effect of n on k^* is always negative. This is because only the older generation supplies capital stock in our model. Given savings, capital stock per worker thus becomes larger when the population of the young is relatively smaller, that is, when n is smaller. The effect of the survival probability q on k^* is, in contrast, ambiguous. It is positive if and only if:

$$1-\alpha > f\alpha\,[q(2+q\beta)-1]. \tag{1.9}$$

When the young have a longer life expectancy, they have more incentive to save in the first period. An increase in q thus has a positive impact on the saving rate of the young. An increase in q, however, reduces the unintended bequest and has a negative impact on the saving. If and only if the former effect dominates the latter, k^* becomes higher as q is larger. The condition (1.9) states that the former effect always dominates the latter when either f, q or α is small enough. We can also verify that even when f, q and α are not small, the former effect dominates for reasonable parameter sets.[6]

Impact on Per Worker Output Growth

We next focus on the linkages of demographic factors with per worker output growth. Since $ln(k_t)=(1/\alpha)\,[ln(y_t)-ln\ A]$, taking the logarithm of (1.7) leads to:

$$ln(y_t)=\alpha\ ln(y_{t-1})+ln\ A+\alpha\ ln\left[\frac{q\beta}{1+q\beta}\,\frac{1-\alpha(1-f)-qf\alpha}{1+n}\right]. \tag{1.10}$$

The average growth rate of per worker output from period 0 to T is thus written as:

$$[ln(y_T)-ln(y_0)]/T=B-\gamma\ ln(y_0)+\delta\ ln\left[\frac{q\beta}{1+q\beta}\,\frac{1-\alpha(1-f)-qf\alpha}{1+n}\right], \tag{1.11}$$

where $B\equiv(\Sigma_{i=0}^{T}\alpha^i/T)\ ln\ A$, $\gamma\equiv(1-\alpha^T)/T$, and $\delta\equiv\alpha\ \Sigma_{i=0}^{T}\alpha^i/T$.

In equation (1.11), γ is the speed of convergence and implies that a country can grow faster when its initial income level is small. The term:

$$ln\left[\frac{q\beta}{1+q\beta}\,\frac{1-\alpha(1-f)-qf\alpha}{1+n}\right]$$

captures the effects of demographic factors on economic growth. It states that the average growth rate of per worker output is higher when the rate of population growth n is smaller, and if (1.9) holds, when the survival probability q is larger. The result is a direct implication from how the steady-state value of per worker capital stock k^* depends on q and n.

Noting that the population ratio of the old to the young, $g\equiv qN_{t-1}/N_t$, is equal to $q/(1+n)$, the term:

$$ln\left[\frac{q\beta}{1+q\beta}\frac{1-\alpha(1-f)-qf\alpha}{1+n}\right]$$

is written as:

$$ln\left\{\frac{g\beta[1-\alpha(1-f)-f\alpha g(1+n)]}{1+(1+n)g\beta}\right\}.$$

Substituting this into (1.11), we obtain:

$$[ln(y_T)-ln(y_0)]/T=B-\gamma\ ln(y_0)+\delta\ ln\left\{\frac{g\beta[1-\alpha(1-f)-f\alpha g(1+n)]}{1+(1+n)g\beta}\right\}. \tag{1.12}$$

Equation (1.12) implies that given the rate of population growth, the average growth rate of per worker output is higher when the ratio of the retirement-age population to the working-age population is larger if (1.9) holds. To the extent that the inequality (1.9) holds, countries with a high proportion of older people therefore tend to achieve high rates of per worker output growth in our model.

Impact on Per Capita Output Growth

Finally, we focus on the linkages of demographic factors with per capita output growth. In our model, per capita output in period t is defined by $z_t\equiv Y_t/(N_t+qN_{t-1})$. Since $y_t\equiv Y_t/N_t$, it holds that $z_t=[N_t/(N_t+qN_{t-1})]\,y_t=[1/(1+g)]\,y_t$, or equivalently, $ln(y_t)=ln(z_t)+ln(1+g)$ for all t. Substituting this into (1.11), the average growth rate of per capita output from period 0 to T is derived as:

$$[ln(z_T)-ln(z_0)]/T=B-\gamma\ ln(z_0)+\delta\ ln\left\{\frac{g\beta[1-\alpha(1-f)-f\alpha g(1+n)]}{1+(1+n)g\beta}\right\}-\gamma\ ln(1+g). \tag{1.13}$$

In particular, a linear approximation leads to:

$$[ln(z_T)-ln(z_0)]/T=\text{constant}-\gamma\cdot ln(z_0)-\chi\cdot n+\lambda\cdot g, \tag{1.14}$$

where:

$$\chi\equiv\delta\ g^*\{f\alpha/[1-\alpha(1-f)-f\alpha\ (1+n^*)g^*]+\beta/[1+(1+n^*)(g^*\beta)]\}$$

and

$$\lambda \equiv (\delta/g^*) - \delta(1+n^*)\{f\alpha/[1-\alpha(1-f) - f\alpha\,(1+n^*)g^*] + \beta/[1+(1+n^*)(g^*\beta)]\} - \gamma/(1+g^*)$$

n^* and g^* are average values of n and g.

Equation (1.14) is a benchmark equation for our empirical studies in the following sections. Since $\chi > 0$, the average growth rate of output becomes higher as n is smaller in (1.14). The linkage of g with per capita output growth is, however, ambiguous in (1.14) even if (1.9) holds. This is because given the growth rate of per worker output, the growth rate of per capita output becomes smaller when g is larger. We can verify that the average growth rate of output becomes higher as g is larger if and only if:

$$\delta/g > \delta(1+n)\{f\alpha/[1-\alpha(1-f) - f\alpha\,(1+n)g] + \beta/[1+(1+n)(g\beta)]\} + \gamma/(1+g). \tag{1.15}$$

When a proportion of the elderly population is large, a significant share of output is consumed by a non-productive segment of the population. To the extent that this negative effect is large, countries with a high elderly dependency rate may have slower rates of economic growth. Countries with a higher proportion of older people, however, tend to have a higher life expectancy and a lower rate of population growth. These demographic factors may thus contribute to capital accumulation per worker and can enhance economic growth. When (1.15) holds, this positive effect dominates the negative effect, and countries with a higher proportion of older people have higher rates of economic growth.

4. SPECIFICATION OF THE BASIC EQUATION

In the following analysis, we examine how well the data will fit to our benchmark equation (1.14) through estimating cross-country growth equations. The cross-country growth equations have a limitation in that the available data series are loose proxies for the relevant variables. The estimations of cross-country growth equations are, however, standard in the previous literature. The standard equation is formulated as follows (see, for example, Barro 1991; Levine and Renelt 1992; and Mankiw et al. 1992):

$$\Delta z/z = \text{constant} - a \cdot ln(z_0) + b \cdot (S/Y) - c \cdot n + d \cdot X, \tag{1.16}$$

where $\Delta z/z$ is the growth rate of per capita income, $ln(z_0)$ is the log of the initial level of per capita income, S/Y is the saving rate (or the investment rate), n is the population growth rate, and X is a factor that may influence the steady-state level of income.

It is well known that the standard equation can be derived by a linear approximation of Solow's growth model (see Mankiw et al. 1992 and Barro and Sala-i-Martin 1995). Our benchmark equation is consistent with the standard equation in that both the initial level of per capita income and the population growth rate have negative impacts on per capita growth rate. However, because it is derived from an overlapping-generations model, our benchmark equation explicitly shows that the ratio of the retirement-age population to the working-age population, g, can be a factor X that influences the steady-state level of income.

A significant difference between the standard equation and our benchmark equation is that the latter does not explicitly include the saving rate (or the investment rate) as an explanatory variable. This is because the saving rate, which is exogenous in Solow's model, is endogenously determined in our model. Our model is, however, too simple to capture all sources of economic growth. We thus include the saving rate (or the investment rate) as an additional explanatory variable in the following estimations.

Since aggregate savings are the sum of positive savings of the young and negative savings of the old in our model, the average saving rate (the ratio of the aggregate savings to total net output) is:

$$(s_t - c_{2\,t})/y_t = [q\beta/(1+q\beta)]\,[(1-\alpha)+(1-q)f\alpha] - [1-(1-q)f]\alpha.$$

This implies that even if the aggregate saving rate is high, the saving rate of the young, that is, $q\beta/(1+q\beta)$, is not necessarily high. The inclusion of the aggregate saving (or investment) rate thus does not rule out the impacts of the saving rate of the young on economic growth in the estimations. One may interpret that the aggregate saving (or investment) rate partly reflects the impacts of non-demographic parameters on economic growth.

5. THE DATA AND BASIC ESTIMATION RESULTS

The Data

The purpose of the following subsections is to estimate equation (1.16) by cross-country data. In the analysis, the data set is supplied by the cross-sectional data set of Penn World Table 6.1.[7] The data is described in detail in Summers and Heston (1988, 1991). We exclude the data of centrally planned countries from the data set. However, unless the data are missing, we run the regression by using the maximum number of countries in the data set. Unless the data are missing, the data set thus includes 140 countries from all regions of the world (see Appendix 1A1 for countries used in the regressions).

The sample periods of the data are 1965–2000. We regress the average growth rate of per capita GDP on the log of per capita real GDP in 1965, the average saving (or investment) rate, and several demographic variables. We also included two regional dummies, that is, dummies of Africa and East Asia. The included explanatory variables are standard in the literature.

A key demographic variable X in the regression is the ratio of the retirement-age population to the working-age population, that is, g in the model. We define the retirement-age population ratio by the ratio of the population above age 65 to those aged 15–64. Given the rate of population growth, our model predicts that it has a positive impact. For other demographic variables, we include the log of population growth rate (and the log of labor population growth rate). See Table 1.1 for the summary statistics and Appendix 1A2 for the detailed definitions.

Table 1.1 Summary statistics

	Mean	Standard error	Min	Max
Real GDP per capita in 1965	4133.60	3809.22	555.74	16644.56
Average growth rate of per capita real GDP, 1965–2000	0.01	0.01	−0.02	0.03
Average investment rate	16.00	8.21	2.29	48.13
Average saving rate	11.23	15.95	−71.30	49.61
Saving rate – Investment rate	−4.77	12.93	−89.78	33.80
Average non-working-age population ratio	0.77	0.17	0.46	1.07
Average child population ratio	0.69	0.21	0.28	1.01
Average of '1–death rate'	0.99	0.01	0.97	1.00
Average retirement-age population ratio	0.09	0.05	0.02	0.25
Average labor population growth rate	0.01	0.01	0.00	0.04
Average life expectancy	60.64	11.57	35.45	77.04
Average population growth rate	2.20	1.18	0.25	9.02

One may argue that the estimations have simultaneous biases because the demographic variables are not necessarily exogenous. It is true that retirement-age population rises and birth rate declines as income level increases. Under such circumstances, there is a clear-cut causality from income level to the demographic factors. However, even in these cases, we can find no definitive causality from economic growth rate to the demographic factors. Most of previous studies thus have assumed that the simultaneous biases, if any, are not serious. We follow this assumption in the following estimations.

Estimation Results

Table 1.2 summarizes our estimation results of equation (1.16) with and without saving (or investment) rate. Regardless of the choice of explanatory variables, the coefficient of the initial per capita income (that is, the log of per capita real GDP in 1980) is significantly negative. This supports the conditional convergence in the world economy.

Concerning the effects of demographic variables, the coefficient of the population growth rate is negative as expected, although it is not significant in some estimates. The coefficient of the retirement-age population ratio, in contrast, shows a significantly positive sign in all cases. This indicates that, given the other variables, a high share of the retirement-age population tends to have strong impacts on economic growth.

When we include the saving (or investment) rate as an explanatory variable, its coefficient is significantly positive. This implies that the saving (or investment) rate provides additional information that explains the sources of economic growth. The inclusion of the saving (or investment) rate, however, did not affect the estimated coefficients of the other explanatory variables, particularly those of demographic variables.

6. ALTERNATIVE ESTIMATION RESULTS

In the previous section, we found that a high share of the retirement-age population tends to have strong impacts on economic growth in the cross-country regressions. The purpose of this section is to examine whether the result is consistent with our model implications in different specifications. Specifically, we estimate alternative equations including some proxies of q in the cross-country regressions. To the extent that our model is correct, it holds that $g \equiv qN_{t-1}/N_t = q/(1+n)$. We would thus obtain the robust results even if we replace the retirement-age population ratio g by proxies of q and n.

As the theory predicts, the retirement-age population ratio g has strong correlations with proxies of q and n in our cross-country data set. For example, when we define q by 'life expectancy', Figure 1.1(a) shows that g has a strong positive correlation with q. Even when we define q by '1 − the death rate', Figure 1.1(b) shows that g has a strong positive correlation with q. In contrast, when we define n by the rate of population growth, Figure 1.2 shows that g has a strong negative correlation with n. Table 1.3 reports the estimation results when we regress g on proxies of q and n. It clearly

Table 1.2 Estimation results of our basic equation

Dependent variable: average growth rate of per capita real GDP, 1965–2000.

	1	2	3	4	5	6	7
Constant	0.0546	0.0403	0.0386	0.0704	0.0448	0.0423	0.0861
	(3.5599)***	(3.0011)***	(2.4299)**	(2.9197)***	(3.1691)***	(2.5873)**	(4.4280)***
Log of real GDP per capita in 1965	−0.0031	−0.0040	−0.0039	−0.0042	−0.0033	−0.0033	−0.0035
	(−2.7402)***	(−3.5538)***	(−3.4687)***	(−3.6508)***	(−2.5468)**	(−2.4581)**	(−2.7420)**
Log of average investment rate	–	0.0058	0.0058	0.0053	–	–	–
		(3.4053)***	(3.3504)***	(3.0487)***			
Log of average saving rate	–	–	–	–	0.0021	0.0022	0.0020
					(3.4281)***	(3.4056)***	(2.8881)***
Log of average retirement-age population ratio	0.0092	0.0069	0.0062	0.0075	0.0062	0.0053	0.0069
	(2.9323)***	(3.6110)***	(2.0268)**	(2.3496)**	(3.4176)***	(1.7590)*	(2.3974)**
Log of average population growth rate	0.0002	–	−0.0005	−0.0037	–	−0.0007	−0.0053
	(0.1750)		(−0.4086)	(−1.4975)		(−0.5415)	(−2.1458)**
Log of average labor population growth rate	–	–	–	0.0051	–	–	0.0075
				(1.6823)*			(2.9443)***
East Asian dummy	0.0117	0.0082	0.0081	0.0081	0.0088	0.0088	0.0086
	(4.8766)***	(3.7130)***	(3.6535)***	(3.9601)***	(3.8128)***	(3.7838)***	(4.1429)***
African dummy	−0.0058	−0.0039	−0.0039	−0.0030	−0.0057	−0.0057	−0.0043
	(−3.7059)***	(−2.3867)**	(−2.3268)**	(−1.7383)*	(−3.1203)***	(−3.0594)***	(−2.3623)**
Number of observations	111	111	111	110	88	88	88
Adjusted R^2	0.4646	0.5535	0.5496	0.5529	0.5089	0.5039	0.5317

Notes:
Heteroskedasticity-consistent *t*-values (White 1980) are reported in parentheses.
*** = significant at 1% level; ** = significant at 5% level; * = significant at 10% level.

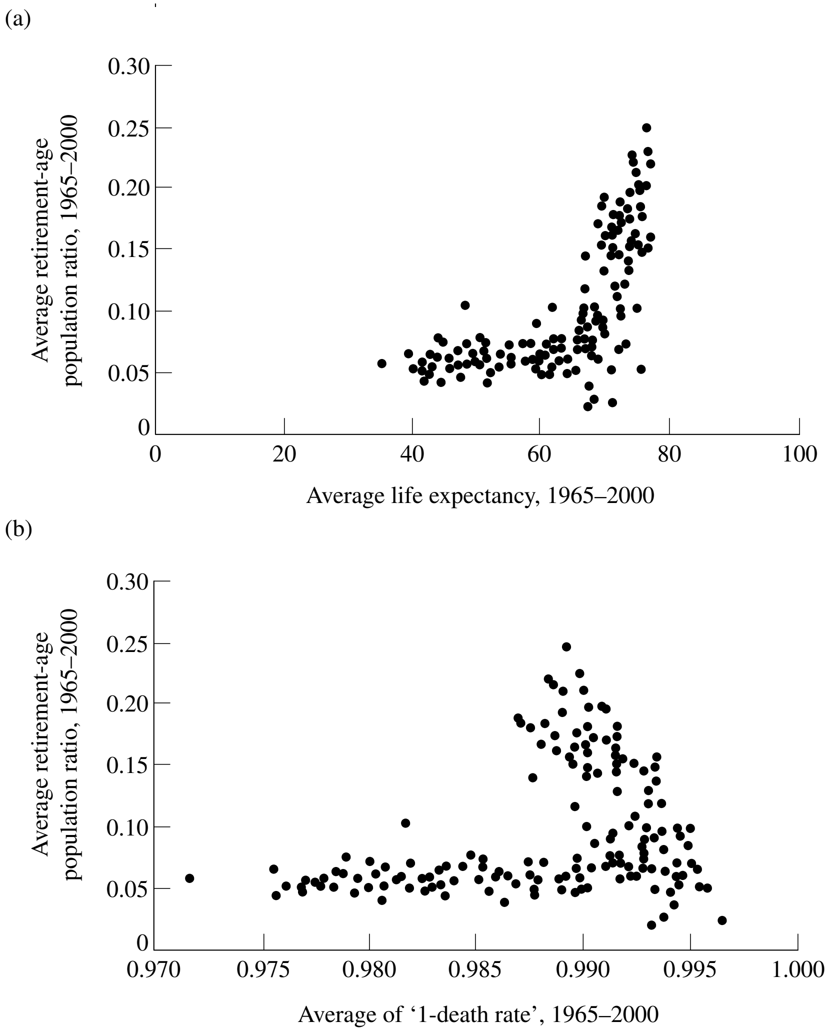


Figure 1.1 Correlations betweeen g *and* q

shows the robust relation among g, q and n. These correlations indicate that the ratio of the retirement-age population to the working-age population can be well approximated by the combination of q and the rate of population growth n.

Table 1.4 summarizes our regression results of our basic equation when we use 'life expectancy' or '1 − the death rate' as an alternative explanatory variable.[8] In the table, the effects of non-demographic explanatory variables

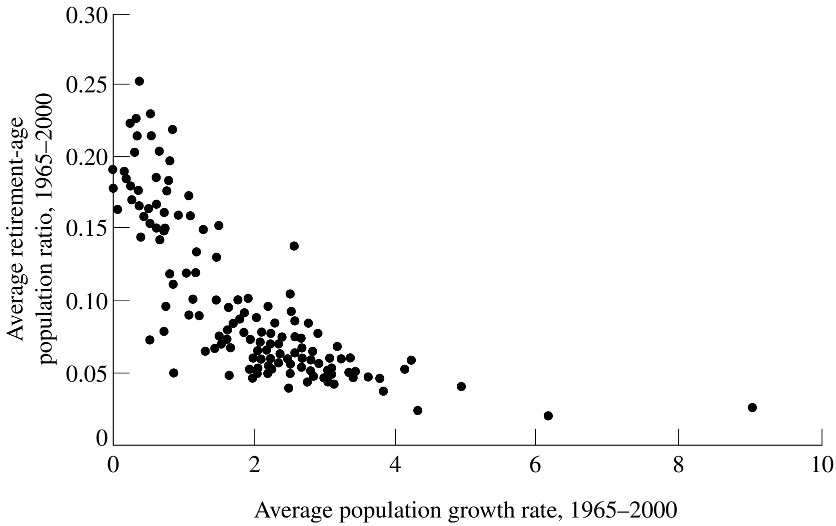


Figure 1.2 Correlations between g *and* n

are essentially the same as those in the last section. In contrast, concerning the effects of q and n, we find a significantly positive sign for the coefficient of q and a significantly negative sign for the coefficient of n. This implies that the positive coefficient of the retirement-age population ratio in Table 1.2 can be attributable to the positive coefficient of q and the negative coefficient of n in Table 1.4.

7. COMPARISONS WITH PREVIOUS STUDIES

So far, we have shown that a rise in the share of persons aged 65 and older in the total population tends to enhance economic growth in many countries. Although the result is consistent with our theoretical model, it may not be consistent with several previous studies that predicted that the declining labor force would bring about a smaller saving rate and slower economic growth. The purpose of this section is to reconcile our results with those in previous literature.

The earlier studies that we mainly discuss in this section are a series of papers by Bloom et al. (Bloom and Williamson 1998 and Bloom et al. 1999). The theoretical parts of these papers were based on Solow's growth model where the saving rates are exogenous. A high proportion of elderly, therefore, have only negative effects on per capita output growth in their

Table 1.3 *Correlations of* g *with* q *and* n

Dependent variable: average retirement-age population ratio, 1965–2000.

	1	2	3	4
Constant	0.0328	−1.4672	−0.0146	−3.9167
	(1.0495)	(−2.5865)**	(−0.6172)	(−11.2561)***
Average population growth rate	−0.0263	−0.0330	–	–
	(−4.3813)***	(−4.5519)***		
Average labor population growth rate	–	–	−6.1283	−7.6518
			(−4.2132)***	(−3.7839)***
Average life expectancy	0.0019	–	0.0028	–
	(5.5082)***		(13.1073)***	
Average of '1–death rate'	–	1.6493	–	4.1382
		(2.9258)***		(11.5703)***
Number of observations	139	139	133	133
Adjusted R^2	0.7077	0.5956	0.7879	0.5990

Notes:
Heteroskedasticity-consistent *t*-values (White 1980) are reported in parentheses.
*** = significant at 1% level; ** = significant at 5% level.

Table 1.4 Estimation results when we used q and n as explanatory variables

Dependent variable: average growth rate of per capita real GDP, 1965–2000

	1	2	3	4
Constant	−0.0853	0.0363	−0.1093	0.0074
	(−4.5102)***	(3.5679)***	(−5.6276)***	(1.0879)
Log of real GDP per capita in 1965	−0.0051	−0.0039	−0.0051	−0.0039
	(−4.5835)***	(−3.7746)***	(−4.6670)***	(−4.0486)***
Log of average investment rate	0.0037	0.0044	0.0036	0.0042
	(2.2974)**	(2.8261)***	(2.2353)**	(2.6891)***
Log of average population growth rate	−0.0025	−0.0046	–	–
	(−3.4356)***	(−4.8664)***		
Log of average labor population growth rate	–	–	−0.0032	−0.0061
			(−3.1853)***	(−4.8603)***
Log of average life expectancy	0.0300	–	0.0321	–
	(4.8673)***		(5.0207)***	
Log of average '1–death rate'	–	0.7329	–	0.9021
		(4.6598)***		(5.0762)***
East Asian dummy	0.0068	0.0065	0.0069	0.0066
	(3.2679)***	(3.0493)***	(3.1312)***	(2.7626)***
African dummy	0.0014	0.0007	0.0011	0.0006
	(0.8736)	(0.3999)	(0.6912)	(0.3183)
Number of observatiions	112	112	110	110
Adjusted R^2	0.6226	0.5994	0.6207	0.6020

Notes:
Heteroskedasticity-consistent *t*-values (White 1980) are reported in parentheses.
*** = significant at 1% level; ** = significant at 5% level.

models. The saving rates are, however, not exogenous in general. In particular, when lifetime horizon is finite, demographic factors such as life expectancies or death rates will have significant impacts on saving rates.

As we discussed in Section 4, the derived growth equation in our overlapping-generations model is similar to that in Solow's model in several aspects. However, because of different assumptions on saving rates, the impacts of population ageing on economic growth in our benchmark equation become completely different from those in the standard equation derived from Solow's model. This implies a fallacy in using the Solow's model to analyse the demographic effects on economic growth.

Concerning the empirical results, our results are no different from those of Bloom et al. In particular, Bloom et al. showed that life expectancy is the most important demographic variable that has a positive impact on economic growth in their cross-country regressions. However, they did not interpret this as evidence that population ageing can have a positive impact on economic growth. Instead, showing that the growth rates of the working-age population have a positive impact on economic growth, they concluded that the decline in the share of the working-age population in total population is harmful for economic growth.

However, interpreting the consequences of a declining share of the working-age population, we need to note that the decline in the working-age population share can be caused either by a rise in the proportion of children or by a rise in the proportion of the elderly. If the decline of the working-age population share is caused mainly by a rise in the proportion of children, the decline in the working-age population share will depress the pace of economic growth. But, if the decline in the working-age population share is caused by a rise in the proportion of the elderly, the decline in the working-age population share does not necessarily slow down rates of economic growth.

To confirm this implication, we estimate our cross-country equation using the ratio of the non-working-age population and the ratio of the child population to the working-age population as explanatory variables. The ratio of the non-working-age population is defined by the ratio of the population below age 14 or over age 65 to those aged 15–64. The ratio of the child population is defined by the ratio of the population below age 14 to those aged 15–64. Table 1.5 reports the estimation results. It clearly shows that both the non-working-age population ratio and the child population ratio have strong negative impacts on economic growth in the regressions.

The results are essentially the same even if we include the retirement age population ratio or the life expectancy as an explanatory variable (see columns 7–11 in Table 1.5). Because the ratio of the child population is highly correlated with the ratio of the retirement age population, we used

Table 1.5 Estimation results with other non-working-age population ratios

Dependent variable: average growth rate of per capita real GDP, 1965–2000

	1	2	3	4	5	6
Constant	0.0236	0.0239	0.0483	0.0211	0.0212	0.0527
	(2.8886)***	(2.9723)***	(2.5115)**	(2.5710)**	(2.5963)**	(2.4328)**
Log of real GDP per capita in 1965	−0.0047	−0.0047	−0.0047	−0.0043	−0.0042	−0.0046
	(−4.0274)***	(−4.0050)***	(−3.8778)***	(−3.6472)***	(−3.4559)***	(−3.5847)***
Log of average investment rate	0.0059	0.0060	0.0055	0.0060	0.0060	0.0054
	(3.3644)***	(3.3647)***	(3.1214)***	(3.4728)***	(3.4665)***	(3.1228)***
Log of average non-working-age population ratio	−0.0191	−0.0181	−0.0174	–	–	–
	(−4.0823)***	(−2.9655)***	(−2.7037)***			
Log of average child population ratio	–	–	–	−0.0097	−0.0091	−0.0115
				(−3.8749)***	(−2.2049)**	(−2.3665)**
Log of average population growth rate	–	−0.0004	−0.0038	–	−0.0003	−0.0037
		(−0.3846)	(−1.6068)		(−0.2641)	(−1.6411)
Log of average labor population growth rate	–	–	0.0046	–	–	0.0056
			(1.4706)			(1.6253)
East Asian dummy	0.0035	0.0037	0.0038	0.0054	0.0056	0.0050
	(1.8972)*	(1.8623)*	(1.9676)*	(2.8240)***	(2.7399)***	(2.6127)**
African dummy	−0.0032	−0.0032	−0.0028	−0.0037	−0.0037	−0.0028
	(−1.9680)*	(−1.9447)*	(−1.5592)	(−2.2202)**	(−2.2115)**	(−1.5035)
Number of observations	112	112	110	111	111	110
Adjusted R^2	0.5634	0.5595	0.5617	0.5529	0.5488	0.5545

	7	8	9	10	11
Constant	0.0502	0.0539	0.0908	−0.0847	−0.1147
	(3.8382)***	(3.3986)***	(3.3815)***	(−4.1916)***	(−5.8821)***
Log of real GDP per capita in 1965	−0.0047	−0.0049	−0.0052	−0.0053	−0.0056
	(−4.2241)***	(−4.1685)***	(−4.3862)***	(−4.5623)***	(−4.8173)***
Log of average investment rate	0.0057	0.0057	0.0051	0.0037	0.0035
	(3.2532)***	(3.1496)***	(2.8027)***	(2.2457)**	(2.1468)**
Log of average retirement-age population ratio	0.0084	0.0096	0.0113	–	–
	(4.3254)***	(3.1380)***	(3.3039)***		
Orthogonal part of logged average child population ratio	−0.0156	−0.0170	−0.0183	−0.0116	−0.0126
	(−1.9486)*	(−2.0834)**	(−2.3069)**	(−1.9720)*	(−2.1700)**
Log of average population growth rate	–	0.0008	−0.0027	−0.0025	–
		(0.7487)	(−1.2288)	(−3.4535)***	
Log of average labor population growth rate	–	–	0.0057	–	−0.0033
			(1.8407)*		(−3.3068)***
Log of average life expectancy	–	–	–	0.0304	0.0341
				(4.5810)***	(5.2244)***
Log of average '1–death rate'	–	–	–	–	–
East Asian dummy	0.0057	0.0056	0.0054	0.0048	0.0047
	(2.5788)**	(2.5658)**	(2.8058)***	(2.2487)**	(2.1082)**
African dummy	−0.0032	−0.0032	−0.0023	0.0016	0.0019
	(−1.8811)*	(−1.8763)*	(−1.2538)	(0.9290)	(1.1109)
Number of observations	111	111	110	111	110
Adjusted R^2	0.5695	0.5664	0.5733	0.6280	0.6318

Notes:
Heteroskedasticity-consistent *t*-values (White 1980) are reported in parentheses.
*** = significant at 1% level; ** = significant at 5% level; * = significant at 10% level.

the orthogonal part of the ratio of the child population in the estimations.[9] The estimation results show that the retirement-age population ratio and life expectancy still have a positive impact on economic growth. In contrast, the ratio of the child population has a strong negative impact on economic growth.[10]

8. THE ROLE OF INTERNATIONAL CAPITAL MOBILITY

In previous sections, we have implicitly assumed a closed-economy model where international capital mobility does not exist. The assumption is, however, restrictive because international capital flows have increased dramatically during the last two decades. The purpose of this section is to investigate how our theoretical and empirical results will change when we allow the role of international capital flows.

In our model, the saving function (1.4) remains the same even when international capital flows exist. However, when capital flows exists, the capital stock is accumulated not only by domestic savings but also by capital inflows. Capital outflows would, on the other hand, slow down the capital stock accumulation. The relationship between the saving of the young and the capital stock thus generally depends on the degree of capital mobility.

Suppose that I/S denotes the share of domestic investment to domestic saving of the young. It then holds that $(1+n)\, k_{t+1} = s_t \cdot (I/S)$. The dynamic equation (1.7) is modified as:

$$k_{t+1} = \frac{q\beta}{1+q\beta} \frac{1-\alpha(1-f)-qf\alpha}{1+n} A\, k_t^{\alpha} \cdot (I/S). \qquad (1.17)$$

Assuming that I/S is exogenous, (1.17) indicates that our basic equation remains the same except that it includes an additional explanatory variable $ln(I/S)$. We thus estimate the following equation:

$$\Delta z/z = \text{constant} - a \cdot ln(z_0) + b \cdot (S/Y) - c \cdot n + d \cdot X + e \cdot ln(I/S), \quad (1.18)$$

where the definitions of variables still follow those in (1.16). There is arbitrariness on the choice of a proxy variable for $ln(I/S)$. We here assume that $ln(I/S)$ is approximated by the average gaps between domestic investment rate and saving rate.

Table 1.6 summarizes our estimation results. Except for the coefficient of $ln(I/S)$, the results are essentially the same as those in previous sections. In particular, the coefficient of the retirement-age population ratio remains significantly positive in all cases, implying that a high share of the retirement-age population tends to have strong impacts on economic growth in many countries. The coefficient of $ln(I/S)$ is, however, significantly positive. This indicates that the impacts of demographic factors on economic growth would be largely reduced if domestic savings were invested in foreign assets.

9. IMPLICATIONS FOR EAST ASIA AND JAPAN

East Asian economies (for example, Japan, Hong Kong, Korea, Taiwan and Singapore) have grown faster than most other developing economies during the past three decades. Several non-demographic factors contributed to the East Asian economic 'miracle'.[11] Recent studies, however, show that demographic variables have also played a large role in East Asia's economic success. A simple international comparison clearly shows that the demographic transition has been more dramatic in East Asia during the past three decades than in any other region. For example, Table 1.7 summarizes the ratios of the working-age population to the total population in various regions. Except for Japan, the ratios in the East Asian economies were not so different from those of other regions in 1970. But, from 1970 to 1990, the ratios rose to significantly higher levels in most of the East Asian economies. East Asia thus appears to have had exceptionally favorable demographic characteristics in the form of high life expectancy and low fertility during this period.

In terms of the ratio of the working-age population, Japan has had favorable demographic characteristics earlier than other East Asian countries. The ratio of the working-age population had already been at a high level in 1970 and remained high until 1995. The demographic forces have by and large sustained Japan's economic growth during the period. The share of elderly people is, however, expected to increase from 17 per cent to 27 per cent in the next 25 years. The labor force, on the other hand, will decline by 0.6 per cent per year. This implies that Japan is going to lose the demographic gift earlier than other East Asian countries.

Given the saving rate of each age group, the demographic change predicts that saving rates in East Asia and Japan would show a long-term decline in the next quarter-century as the population ages.[12] However, how serious the problem will be crucially depends on how domestic savings sustain the potential growth rate of the economy. Even when the population ages, Japan's saving rates may remain high because the saving rates of the working-age population will increase.[13] Our cross-country evidence supported this possibility.

The result in the last section, however, suggests that capital outflows

Table 1.6 Estimation results with IS *gap*

Dependent variable: average growth rate of per capita real GDP, 1965–2000

	1	2	3	4	5	6	7	8	9
Constant	0.0479	0.0466	0.0812	−0.0831	0.0462	−0.1042	0.0180	−0.1037	0.0277
	(3.6512)***	(2.9816)***	(3.2388)***	(−4.3499)***	(4.1602)***	(−5.1626)***	(2.0361)**	(−3.8116)***	(1.5007)
Log of real GDP per capita in 1965	−0.0040	−0.0040	−0.0043	−0.0054	−0.0043	−0.0055	−0.0042	−0.0055	−0.0043
	(−3.1094)***	(−3.0343)***	(−3.0348)***	(−3.9143)***	(−3.3064)***	(−3.9272)***	(−3.4342)***	(−3.9227)***	(−3.3941)***
Log of average saving rate	0.0004	0.0004	0.0004	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003
	(3.2187)***	(3.1887)***	(2.9069)***	(2.2865)**	(2.9570)***	(2.1522)**	(2.7610)***	(2.1473)**	(2.7353)***
Saving rate – investment rate	−0.0004	−0.0004	−0.0004	−0.0002	−0.0003	−0.0002	−0.0003	−0.0002	−0.0003
	(−3.4005)***	(−3.3518)***	(−3.0304)***	(−2.0551)**	(−2.9555)***	(−1.9610)*	(−2.7518)***	(−1.9566)*	(−2.7487)***
Log of average retirement-age population ratio	0.0062	0.0057	0.0071	–	–	–	–	–	–
	(3.1660)***	(1.7752)*	(2.1390)**						
Log of average population growth rate	–	−0.0004	−0.0039	−0.0022	−0.0044	–	–	−0.0000	−0.0016
		(−0.2795)	(−1.5434)	(−3.1880)***	(−4.8198)***			(−0.0232)	(−0.6945)
Log of average labor population growth rate	–	–	0.0056	–	–	−0.0029	−0.0059	−0.0029	−0.0040
			(1.8012)*			(−2.9497)***	(−4.7254)***	(−1.0547)	(−1.2696)
Log of average life expectancy	–	–	–	0.0315	–	0.0331	–	0.0331	–
				(5.1506)***		(5.1947)***		(5.1743)***	

Log of average ‘1–death rate’	–	–	–	–	0.7836 (5.0089)***	–	0.9350 (5.3393)***	–	0.8869 (5.0090)***
East Asian dummy	0.0062 (2.3579)**	0.0062 (2.3214)**	0.0064 (2.5283)**	0.0053 (2.1434)**	0.0047 (1.8115)*	0.0055 (2.1157)**	0.0048 (1.7267)*	0.0055 (2.1132)**	0.0049 (1.7993)*
African dummy	−0.0046 (−2.8877)***	−0.0046 (−2.8367)***	−0.0036 (−2.1170)**	0.0013 (0.7813)	0.0006 (0.3474)	0.0010 (0.5649)	0.0004 (0.2267)	0.0010 (0.5618)	0.0005 (0.2611)
Number of observations	111	111	110	112	112	110	110	110	110
Adjusted R^2	0.5336	0.5293	0.5353	0.6158	0.5956	0.6131	0.5983	0.6092	0.5962

Notes:
Heteroskedasticity-consistent *t*-values (White 1980) are reported in parentheses.
*** = significant at 1% level; ** = significant at 5% level; * = significant at 10% level.

Table 1.7 Ratios of the working-age population to total population

Region	Years	Ratio
Sub-Saharan Africa	1970	0.53
	1990	0.52
Latin America and the Caribbean	1970	0.52
	1990	0.57
Asia	1970	0.55
	1990	0.62
Southeast Asia	1970	0.53
	1990	0.59
South Asia	1970	0.53
	1990	0.57
East Asia	1970	0.59
	1990	0.69
Taiwan	1970	0.57
	1990	0.68
	1995	0.69
Korea	1970	0.57
	1990	0.68
	1995	0.70
Hong Kong	1970	0.62
	1990	0.70
	1995	0.71
Singapore	1970	0.56
	1990	0.70
	1995	0.70
Japan	1970	0.69
	1990	0.69
	1995	0.70

Note: The total working-age population is defined by age 15 to 64.

Source: The data is based on World Bank data.

would reduce the impact of the increased savings. Large fiscal deficits that shift private savings to the public sector would be another negative symptom to sustain capital accumulation in the future. During the past two decades, Japan has experienced large capital outflows and large fiscal deficits accumulated throughout the 1990s. These symptoms probably imply that our results cannot predict an optimistic future for the Japanese economy.

10. CONCLUDING REMARKS

Because people's economic behavior varies at different stage of life, changes in a nation's age structure can have significant effects on its economic growth. If most of a country's population falls within the working age group, the added productivity of this group can produce a 'demographic gift' of economic growth. In contrast, countries with a high proportion of children are likely to devote a large part of resources to their care, which tends to depress the pace of economic growth. The main purpose of this chapter was to investigate whether a country can sustain economic growth when a large proportion of a country's population consists of the elderly. When the proportion of the elderly is large in the total population, a large share of resources is needed by a relatively less productive segment of the population. To the extent that this negative effect is large, countries with a high elderly dependency rate can thus have slower rates of economic growth. Countries with a high proportion of elderly people, however, tend to have a high life expectancy and a low rate of population growth. These demographic factors increase the saving rate of the working-age population and may enhance economic growth. It is therefore not clear whether countries with a high proportion of the elderly tend to have smaller rates of economic growth or not.

The first part of this chapter theoretically formulates this intuition by an overlapping-generations model. The second part then empirically examined it by cross-country regressions. Estimating standard cross-country growth equations, we observed significantly positive linkages between shares of the elderly and rates of economic growth. The results suggest that a country may sustain economic growth even after the elderly dependency rate rises. Needless to say, the cross-country evidence is not sufficient to present a satisfactory prediction for the future in a specific country because each country has different economic and institutional environments. In order to derive a more definitive answer to a specific country, it is highly desirable to investigate more detailed data for each country. Although the analysis is beyond the scope of this chapter, this is an important topic for our research project.

APPENDIX 1A1 COUNTRIES USED IN CROSS-COUNTRY REGRESSIONS: 140 COUNTRIES

Afghanistan
Algeria
Angola
Argentina
Australia
Austria
Bahamas
Bahrain
Bangladesh
Barbados
Belgium
Belize
Benin
Bhutan
Bolivia
Botswana
Brazil
Brunei
Burkina Faso
Burundi
Cameroon
Canada
Cape Verde
Central African Republic
Chad
Chile
Colombia
Comoros
Congo, Dem. Rep.
Congo, Rep.
Costa Rica
Côte d'Ivoire
Cyprus
Denmark
Djibouti
Dominican Republic
Ecuador
Egypt, Arab Rep.
El Salvador
Equatorial Guinea
Ethiopia
Fiji
Finland
France
Gabon
Gambia
Germany
Ghana
Greece
Guam
Guatemala
Guinea
Guinea-Bissau
Guyana
Haiti
Honduras
Hong Kong, China
Iceland
India
Indonesia
Iran, Islamic Rep.
Iraq
Ireland
Israel
Italy
Jamaica
Japan
Jordan
Kenya
Korea, Rep.
Kuwait
Lebanon
Lesotho
Liberia
Libya
Luxembourg
Macao, China
Madagascar
Malawi
Malaysia
Maldives
Mali
Malta
Mauritania
Mauritius
Mexico
Morocco
Mozambique
Namibia
Nepal
Netherlands
New Caledonia
New Zealand
Nicaragua
Niger
Nigeria
Norway
Oman
Pakistan
Panama
Papua New Guinea
Paraguay
Peru
Philippines
Portugal
Puerto Rico
Qatar
Rwanda
Saudi Arabia
Senegal
Seychelles

Sierra Leone
Singapore
Solomon Islands
Somalia
South Africa
Spain
Sri Lanka
St. Lucia
Sudan
Suriname
Swaziland
Sweden
Switzerland
Syrian Arab Republic
Tanzania
Thailand
Togo
Trinidad and Tobago
Tunisia
Turkey
Uganda
United Arab Emirates
United Kingdom
United States
Uruguay
Vanuatu
Venezuela, Bolivarian Republic
Zambia
Zimbabwe

APPENDIX 1A2 THE DEFINITIONS OF VARIABLES USED IN REGRESSIONS

1. Real per capita GDP = Real per capita GDP terms of trade adjustment (1996 prices).
   Source: Penn World Table 6.1 (*RGDPTT*).
   The *RGDPTT* variable is gross domestic income and follows the recommended method in the UN System of National Accounts. This revised procedure is also consistent with the current and past treatment of the net foreign balance in the Penn World Table. *RGDPTT* is the 1996 international price value of domestic absorption of a country in a given year plus current exports minus current imports deflated by the deflator and the 1996 purchasing power parity of domestic absorption.

2. Log of average domestic investment rate = Log of average *CI* from 1965 to 2000.
   Source: Penn World Table 6.1 (*CI*).
   The *CI* variable is investment share of *CGDP*.
   The *CGDP* variable is real GDP per capita.

3. Log of average domestic saving rate = Log of average *CSAVE* from 1965 to 2000.
   Source: Penn World Table 6.1 (*CSAVE*).
   The *CSAVE* variable is current savings. This variable is defined as the percentage share of current savings to GDP and is derived by subtracting *CC* and *CG* from 100.

The *CC* variable is consumption share of *CGDP*.
The *CG* variable is government share of *CGDP*.

4. Saving rate − Investment rate = Average of *CSAVE-CI* from 1965 to 2000.
   Source: Penn World Table 6.1 (*CSAVE*, *CI*).

5. Log of average retirement-age population ratio = Log of average of (population aged 65 and above, total / population aged 15–64, total) from 1965 to 2000.
   Source: World Development Indicators 2002 (CD-ROM).
   Population aged 65 and above, total: Total population 65 years of age or older. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship – except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of the country of origin.
   Population ages 15–64, total: Total population between the ages 15 to 64 is the number of people who could potentially be economically active. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship – except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of the country of origin.

6. Average population growth rate = Average population growth (annual %) from 1965 to 2000,
   Source: World Development Indicators 2002 (CD-ROM).
   Population growth (annual %): Annual population growth rate. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship – except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of the country of origin.

7. Average labor population growth rate = Average labor population growth rate (annual) induced from labor force total in 1965 and 2000.
   Source: World Development Indicators 2002 (CD-ROM).
   Labor force, total: Total labor force comprises people who meet the International Labour Organization definition of the economically active population: all people who supply labor for the production of goods and services during a specified period. It includes both the employed and the unemployed. While national practices vary in the

treatment of such groups as the armed forces and seasonal or part-time workers, in general the labor force includes the armed forces, the unemployed, and first-time job-seekers, but excludes homemakers and other unpaid caregivers and workers in the informal sector.

8. Average life expectancy = Average life expectancy at birth, total (years) from 1965 to 2000.
Source: World Development Indicators 2002 (CD-ROM).
Life expectancy at birth, total (years): Life expectancy at birth indicates the number of years a newborn infant would live if prevailing patterns of mortality at the time of its birth were to stay the same throughout its life.

9. Average of '1-death rate' 1965–2000 = Average [1-death rate, crude (per 1000 people)/1000] from 1965 to 2000.
Source: World Development Indicators 2002 (CD-ROM).
Death rate, crude (per 1000 people): Meaning of survival rates. Crude death rate indicates the number of deaths occurring during the year, per 1000 population estimated at mid-year.

10. Log of average child population ratio 1965–2000 = Log of average ratio of population below aged 14 to those aged 15–64 from 1965 to 2000.
Child population ratio = (Population aged 0–14, total/Population aged 15–64, total).
Source: World Development Indicators 2002 (CD-ROM).
Population aged 0–14, total: Total population between the ages 0 and 14. Population is based on the de facto definition of population, which counts all residents regardless of legal status or citizenship – except for refugees not permanently settled in the country of asylum, who are generally considered part of the population of the country of origin.

11. African countries used for the African dummy: Algeria, Angola, Benin, Botswana, Burkina Faso, Burundi, Cameroon, Cape Verde, Central African Republic, Chad, Comoros, Congo, Dem. Rep., Congo, Rep., Côte d'Ivoire, Djibouti, Egypt, Arab Rep., Equatorial Guinea, Ethiopia, Gabon, Gambia, Ghana, Guinea, Guinea-Bissau, Kenya, Lesotho, Liberia, Libya, Madagascar, Malawi, Mali, Mauritania, Mauritius, Morocco, Mozambique, Namibia, Niger, Nigeria, Rwanda, Senegal, Seychelles, Sierra Leone, Somalia, South Africa, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia, Zimbabwe.

12. East Asian countries used for the East Asian dummy: Brunei, Hong Kong, China, Indonesia, Japan, Korea, Rep., Macao, China, Malaysia, Philippines, Singapore, Thailand.

NOTES

* An earlier version of this chapter was presented at the International Forum of Collaboration Projects sponsored by the Economic and Social Research Institute, Cabinet Office, Government of Japan. We would like to thank Ales Cerny, Edward Prescott and other participants for their helpful comments.

1. These studies include Fry and Mason (1982), Bloom and Freeman (1988), Mason (1988), Masson (1990), Webb and Zia (1990), Taylor (1995), Kelley and Schmidt (1996), Higgins (1998), and Lindh and Malmberg (1999). Most of the studies investigate the role of savings in explaining linkages between demographic change and economic growth. There are also several case studies for specific countries: McMillan and Baesel (1990) and Fair and Dominguez (1991) for the United States and Malmberg (1994) for Sweden. Recent work by Loayza et al. (2000) extensively studied linkages between demographic change and savings.
2. In a different context, Krugman (1994) also predicted that economic growth in East Asia would likely slow in the future because it had been supported by transitorily large factor inputs.
3. In explaining high saving rates in Japan, Horioka (1991a) attributed quite a lot of explanatory power to variations in the age structure.
4. See, for examples, several papers in Hurd and Yashiro (1997).
5. More recently, Bryant et al. (2002) and McKibbin and Nguyen (2002) derived similar conclusions in their simulation models.
6. For example, suppose that $\alpha = 1/3$. Then, we can show that the inequality (1.9) holds for any combination of β, q, and f.
7. In the previous version, we used the data set of World Development Indicators (World Bank) and obtained essentially the same results.
8. In the table, we only reported the case where the investment rate is included as an explanatory variable. The results, however, did not change in other specifications.
9. The orthogonal part is obtained from residuals through regressing it on the ratio of the retirement-age population.

10 In a recent empirical study, Loayza et al. (2000) showed cross-country evidence that both the ratio of the child population and the ratio of the retirement-age population have significantly negative impacts on saving rates. The results, however, seem to be derived because growth rates of income were included in one of the explanatory variables. Even in their data set, we find that the ratio of the retirement-age population is positively correlated with saving rates, while the ratio of the child population is negatively correlated with saving rates (see Table 1.3).

11. These factors include outward-looking development strategies, high domestic saving rates, strong inflows of foreign direct investment, technological ‘catch-up’, relatively low income inequality, a stable macroeconomic environment, a market-friendly policy environment and so on. See, for example, World Bank (1993) and Easterly (1995).
12. See, for example, Horioka (1991b).
13. In their simulation model, Faruqee and Mühleinsen (2001) reached the same conclusion.

REFERENCES

Barro, R.J. (1991), 'Economic growth in a cross section of countries', *Quarterly Journal of Economics*, **106**, 407–43.

Barro, R.J. and X. Sala-i-Martin (1995), *Economic Growth*, New York: McGraw-Hill.

Bloom, D.E., D. Canning and P.N. Malaney (1999), 'Demographic change and economic growth in Asia', Center for International Development (CID) Working Paper no. 15, Harvard University.

Bloom, D.E. and R.B. Freeman (1988), 'Economic development and the timing and componets of population growth', *Journal of Policy Modeling*, **10**, 57–81.

Bloom, D.E. and J.G. Williamson (1998), 'Demography transitions and economic miracles in emerging Asia', *World Bank Economic Review*, **12**, 419–55.

Bryant, R.C., H. Faruqee and D. Velculescu (2002), 'Youth dependency and demographic change: implications for the global economy', paper presented at the International Forum of Collaboration Projects, Tokyo, Spring.

Coale , A.J. and E. Hoover (1958), *Population Growth and Economic Development in Low-income Countries*, Princeton, NJ: Princeton University Press.

Diamond, P. (1965), 'National debt in a neoclassical growth model', *American Economic Review*, **55**, 1126–50.

Easterly, W. (1995), 'Explaining miracles: growth regressions meet the Gang of Four', in T. Ito and A.O. Krueger (eds), *Growth Theories in Light of East Asian Experience*, Chicago: University of Chicago Press, 267–90.

Ehrlich, I. and F.T. Lui (1991), 'Intergenerational trade, longevity, and economic growth', *Journal of Political Economy*, **99**, 1029–59.

Fair, R.C. and K.M. Dominguez (1991), 'Effects of changing U.S. age distribution on macroeconomic equations', *American Economic Review*, **81**, 1276–94.

Faruqee, H. and M. Mühleinsen (2001), 'Population aging in Japan: demographic shock and fiscal sustainability', IMF Working Paper WP/01/40, International Monetary Fund, Washington, DC.

Fry, M. and A. Mason (1982), 'The variable rate-of-growth effect in the life-cycle saving model: children, capital inflows, interest and growth in a new specification of the life-cycle model applied to seven Asian developing countries', *Economic Inquiry*, **20**, 426–42.

Higgins, M. (1998), 'The demographic determinants of savings, investment and international capital flows', *International Economic Review*, **39**, 343–69.

Higgins, M. and J.G. Williamson (1996), 'Asian demography and foreign capital dependence', NBER Working Paper no. 5560, National Bureau of Economic Research, Cambridge, MA.

Higgins, M. and J.G. Williamson (1997), 'Age structure dynamics in Asia and dependence of foreign capital', *Population and Development Review*, **23**, 261–93.

Horioka, C.Y. (1991a), 'The determinants of Japan's saving rate: the impact of the age structure of the population and other factors', *Economic Studies Quarterly*, **42**, 237–53.

Horioka, C.Y. (1991b), 'Future trends in Japan's saving rate and the implications thereof for Japan's external imbalance', *Japan and the World Economy*, **3**, 307–30.

Hurd, M.-D. and N. Yashiro, (eds) (1997), *The Economic Effects of Aging in the United States and Japan*, Chicago: University of Chicago Press.

Kalemli-Ozcan, S., H.E. Ryder and D.N. Weil (2000), 'Mortality decline, human

capital investment, and economic growth', *Journal of Development Economics*, **62**, 1–23.
Kelley, A.C. and R.M. Schmidt (1996), 'Saving, dependency and development', *Journal of Population Economics*, **9**, 365–86.
Krugman, P. (1994), 'The myth of Asia's miracle', *Foreign Affairs*, **73**, 62–78.
Levine, R. and D. Renelt (1992), 'A sensitivity analysis of cross-country growth regressions', *American Economic Review*, **82**, 942–63.
Lindh, T. and B. Malmberg (1999), 'Age structure effects and growth in the OECD, 1950–1990', *Journal of Population Economics*, **12**, 431–49.
Loayza, N., K. Schmidt-Hebbel and L. Servén (2000), 'What derives private saving across the world', *Review of Economics and Statistics*, **82**, 165–81.
Malmberg, B. (1994), 'Age structure effects on economic growth – Swedish evidence', *Scandinavian Economic History Review*, **42**, 279–95.
Mankiw, N.G., D. Romer and D.N. Weil (1992), 'A contribution to the empirics of economic growth', *Quarterly Journal of Economics*, **107**, 407–37.
Mason, A. (1988), 'Saving, economic growth, and demographic change', *Population and Development Review*, **14**, 113–44.
Masson, P. (1990), 'Long-term Macroeconomic Effects of Aging Populations', *Finance and Development*, **27**, 6–9.
McKibbin, W.J. and J. Nguyen (2002), 'Demographic change in Japan: projected impacts in an empirical model with and without children', paper presented at the International Forum of Collaboration Projects, Tokyo, Spring.
McMillan, H.M. and J.B. Baesel (1990), 'The macroeconomic impact of the baby boom generation', *Journal of Macroeconomics*, **12**, 167–95.
Solow, R. (1956), 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, **70**, 65–94.
Summers, R. and A. Heston (1988), 'A new set of international comparisons of real product and price levels estimates for 130 countries, 1950–1985', *Review of Income and Wealth*, **34**, 1–25.
Summers, R. and A. Heston (1991), 'The Penn World Table (Mark 5): an expanded set of international comparisons, 1950–1988', *Quarterly Journal of Economics*, **106**, 327–68.
Taylor, A. (1995), 'Debt, dependence and the demographic transition: Latin America into the next century', *World Development*, **23**, 869–79.
Webb, S. and H. Zia (1990), 'Lower birth rates = higher saving in LDCs', *Finance and Development*, **27**, 12–14.
White, H. (1980), 'A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity', *Econometrica*, **48**, 817–38.
World Bank (1993), *The East Asian Miracle*, New York: Oxford University Press.

2. The 1990s in Japan: a lost decade*

Fumio Hayashi and Edward C. Prescott

1. INTRODUCTION

The performance of the Japanese economy in the 1990s was less than stellar. The average annual growth rate of per capita GDP was 0.5 percent in the 1991–2000 period. The comparable figure for the United States was 2.6 percent. Japan in the last decade, after steady catch-up for 35 years, not only stopped catching up but lost ground relative to the industrial leader. The question is why.

A number of hypotheses have emerged: inadequate fiscal policy, the liquidity trap, depressed investment due to overinvestment during the 'bubble' period of the late 1980s and early 1990s, and problems with financial intermediation. These hypotheses, while possibly relevant for business cycles, do not seem capable of accounting for the chronic slump seen ever since the early 1990s. This chapter offers a new account of the 'lost decade', based on the neoclassical growth model.

Two developments are important for the Japanese economy in the 1990s. First and most important is the fall in the growth rate of total factor productivity (TFP). This had the consequence of reducing the slope of the steady-state growth path and increasing the steady-state capital–output ratio. If this were the only development, investment share and labor supply would decrease to their new lower steady-state values during the transition. But the drop in the *rate* of productivity growth alone cannot account for the near-zero output growth in the 1990s.

The second development is the reduction of the workweek length (average hours worked per week) from 44 hours to 40 hours between 1988 and 1993, brought about by the 1988 revision of the Labor Standards Law. In the most standard growth model, where aggregate hours (average hours worked times employment) enter the utility function of the stand-in consumer, a decline in workweek length does not affect the steady-state growth path because the decline is offset by an increase in employment. However, in our specification of the growth model, the workweek length and employment enter the utility function separately, so that a shortening of the workweek shifts the *level* of the steady-state growth path down. If the only change were a reduction in workweek length, the economy would converge to a lower steady-state growth path subsequent to the reduction in the workweek length.

We determine the consequence of these two factors for the behavior of the Japanese economy in the 1990s. To do this we calibrate our growth model to pre-1990 data and use the model to predict the path of the economy in the 1990s, treating TFP as exogenous and treating the workweek length as endogenous subsequent to 1993. The lost decade of growth is what the model predicts. Also predicted is the increase in the capital–output ratio and the fall in the return on capital that occurred through the 1990s. The only puzzle is why the TFP growth was so low subsequent to 1991. We discuss possible reasons for this decline in the concluding section.

In Section 2, we start with a brief catalogue of some of the facts about the lost decade. We then proceed to examine the Japanese economy through the perspective of growth theory in Section 3. We use this model economy to predict what will happen in the 1990s and beyond, taking the paths of productive efficiency, workweek lengths, capital tax rate and the output share of government purchases as exogenous.

Growth theory gives no role to frictions in financial intermediation. To many this may appear a serious omission. It is natural to suspect that the collapse of bank loans that took place throughout the 1990s must have something to do with the output slump in the same decade. There is an emerging literature about Japan that asks (a) whether the decline in bank loans was a 'credit crunch' – namely, a decline due to supply factors (such as the Bank for International Settlements (BIS) capital ratio imposed on banks), and (b) if so, whether it depressed output by constraining investment.[1] In Section 4, we present evidence from various sources that the answer to the first question is probably 'yes', but the answer to the second question is 'no'. That is, despite the collapse of bank loans, firms found ways to finance investment. This justifies our neglect of financial factors in accounting for the lost decade. Section 5 contains concluding remarks.

2. THE JAPANESE ECONOMY, 1984–2000

We begin with an examination of the NIA (National Income Accounts) data for the 1984–2000 period and report the facts that are most germane to real growth theory, which abstracts from monetary and financial factors. In this section, we shall determine the importance of the TFP behavior and the reduction in the workweek length on the behavior of the Japanese economy in the 1990s.

Poor Performance in the 1990s

Figure 2.1 documents Japan's prolonged slump in the 1990s. The figure graphs the real GNP per adult (aged 20–69), detrended at 2 percent (which has been the long-run growth rate for the leader country over the past century) and normalized to 100 for 1990.[2] The performance of the

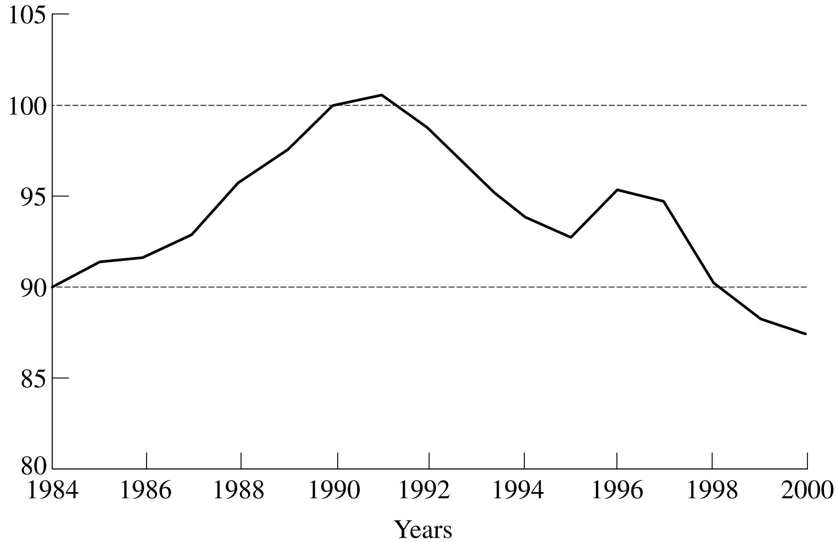


Figure 2.1 Detrended real GNP per working-age population

economy was very good in the 1980s, growing at a much higher rate than the benchmark 2 percent, and looking as if poised to catch up with the United States. However, this trend reversed itself subsequent to 1991, and by 2000 the Japanese per adult GNP is less than 90 percent of what it would have been had it kept growing at 2 percent since 1991. Part of this slowdown is due to a decline in TFP growth. Over the 1983–91 period, TFP grew at a more than respectable rate of 2.4 percent.[3] It fell to an average of 0.2 percent for 1991–2000.

Workweek Falls in the 1988–93 Period

An important policy change occurred at the end of the 1980s. The workweek length declined from 44 hours in 1988 to 40 hours in 1993, as depicted in Figure 2.2. This decline was by government fiat. For the first time in 40 years, there was a major revision of the Labor Standards Law in 1988, which stipulated a gradual reduction in the statutory workweek[4] from 48

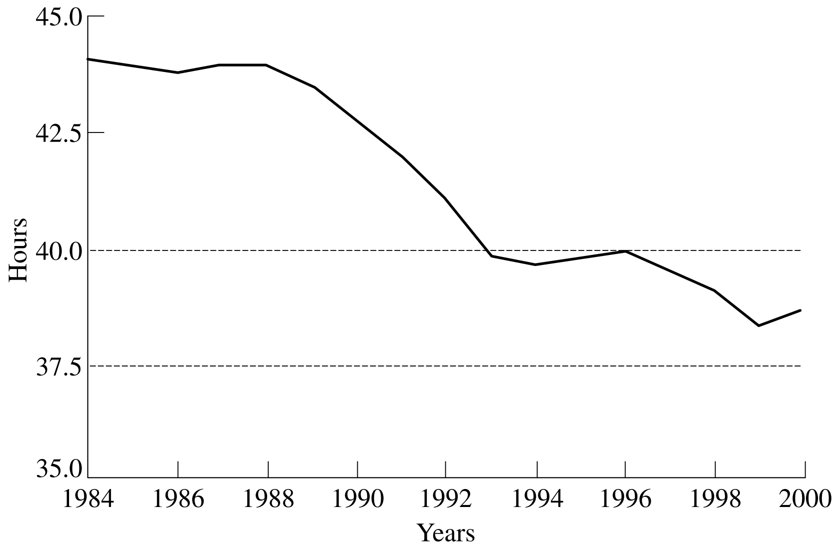


Figure 2.2 Length of workweek

hours down to 40 hours (six down to five workdays per week), to be phased in over several years. The number of national holidays increased by three during this period. Government offices were closed on Saturdays every other week beginning in 1989, and since 1992 have been closed every Saturday. Financial institutions have been closed every Saturday since 1989. A new temporary law was introduced in 1992 to bring about further reduction in hours worked. The 1998 revision of the Labor Standards Law added one day to paid vacation. It appears that the government's drive to reduce workweek had a lot of public support, judging from newspaper accounts.

Capital Deepens as the Rate of Return Declines in the 1990s

Figure 2.3 plots the nongovernment capital–output ratio. An accounting convention we follow throughout the chapter is that all government purchases are expensed (that is, treated as consumption) and that the current account balance (the sum of net exports and net factor income from abroad) is included as investment. Therefore, the capital stock excludes government capital but includes claims on the rest of the world (foreign capital). Following theory, we include inventory stocks as part of the capital stock. Looking at Figure 2.3, we note that there was a significant capital deepening, with the capital–output ratio increasing by nearly 30

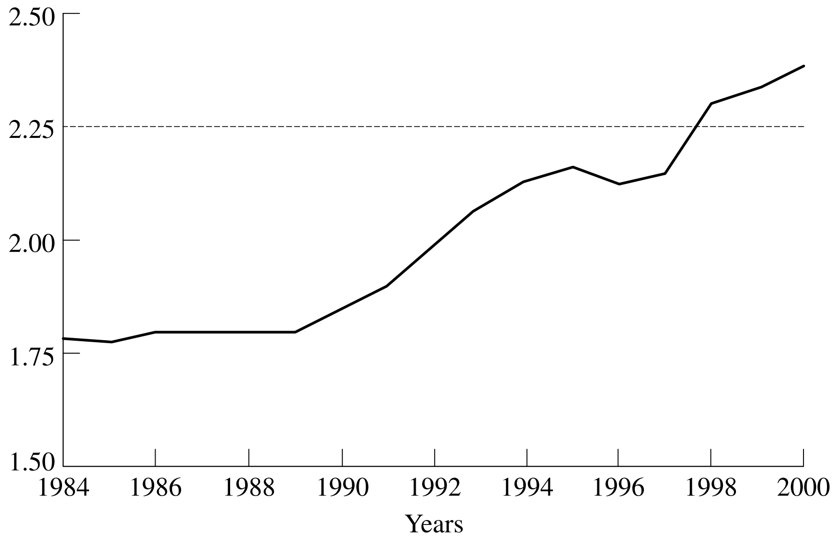


Figure 2.3 Capital–output ratio

percent, from 1.86 at the beginning of 1990 to 2.39 in 2000. (If the capital stock excludes foreign capital, the ratio increases from 1.67 in 1990 to 1.98 in 2000.) Associated with this capital deepening, there was a decline in the after-tax (and net) return on capital, depicted in Figure 2.4, from 6.1 percent in the late 1980s to 4.2 percent in the late 1990s.

Both these rate of return figures are too high because part of the return includes the return on land. To get a better idea of the levels of return as opposed to just the change in returns, we examine returns in the non-land intensive sectors, namely the corporate and foreign sectors. The decline in after-tax profits divided by capital stocks in these sectors is from 5.3 percent to 2.1 percent. This leads us to the assessment that the after-tax return on capital declined over three percentage points between 1990 and 2000: from over 5 percent to about 2 percent.

Government Share Increases and Investment Share Decreases in the 1990s

Figure 2.5 shows that the composition of output changed in the 1990s. The government's share of output increased from an average share of 13.7 percent in the 1984–90 period to 15.2 percent in the 1994–2000 period. Another change is the decline in private investment share from 27.6 percent to 24.3 percent in these periods. Most of the decline in investment occurred in the domestic investment component, not in the current account: the output share

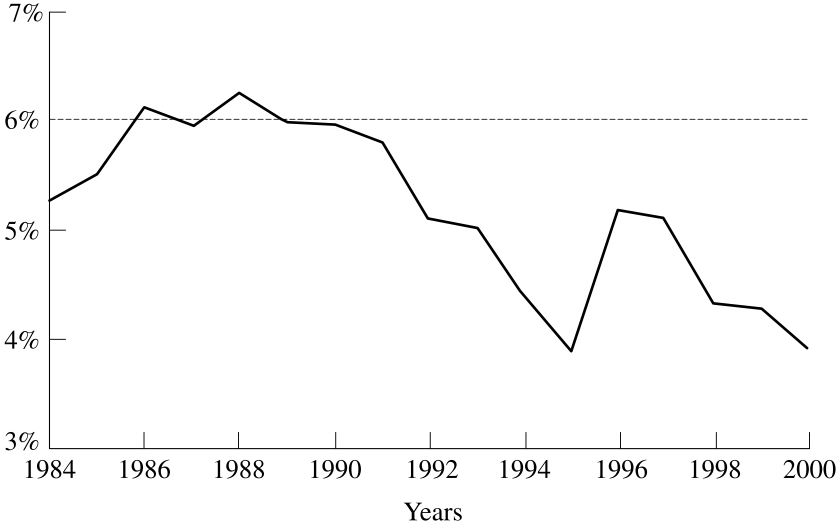


Figure 2.4 After-tax rate of return

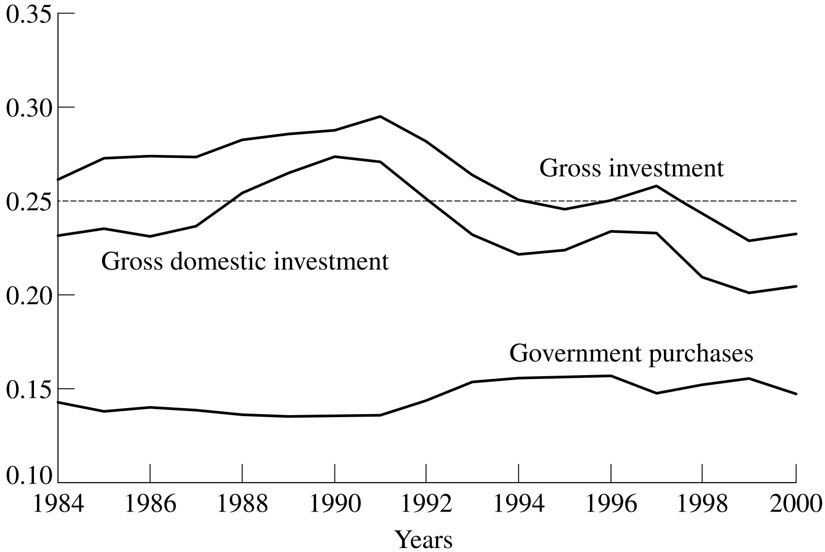


Figure 2.5 Government purchases and investment as a share of output

of domestic investment declined by 3 percentage points, from 24.6 percent to 21.7 percent. The decline in the late 1990s is rather substantial.

3. JAPANESE ECONOMY FROM THE GROWTH THEORY PERSPECTIVE

In using growth theory to view the Japanese economy in the 1990s, we are using a theory that students of business cycles use to study business cycles and students of public finance use to evaluate tax policies. The standard growth model, however, must be modified in one important way to take into account the consequences of a policy change that led to a reduction in the average workweek in Japan in the 1988–93 period. Taking as given the fall in the workweek length, the fall in productivity growth and the increase in the output share of government purchases in the 1990s, we use the theory to predict the path of the Japanese economy after 1990.

The Growth Model

Technology

The aggregate production function is:

$$Y = AK^{\theta}(h\,E)^{1-\theta}, \tag{2.1}$$

where Y is aggregate output, A is TFP, K is aggregate capital, E is aggregate employment and h is hours per employee.

Growth accounting

Having specified the aggregate production function, we can go back to the data on the Japanese economy and perform growth accounting. Our growth accounting, involving the capital–output ratio instead of the capital stock, is equivalent to, but differs in appearance from, the usual growth accounting. Let N be the working-age population and define:

$$y \equiv Y/N,\; e \equiv E/N,\; x \equiv K/Y. \tag{2.2}$$

Using these definitions on (2.1) and by simple algebra, we obtain:

$$y = A^{1/(1-\theta)}\,h\,e\,x^{\theta/(1-\theta)}. \tag{2.3}$$

That is, output per adult y can be decomposed into four factors: the TFP factor $A^{1/(1-\theta)}$, the workweek factor h, the employment rate factor e and the capital intensity factor $x^{\theta/(1-\theta)}$. Our growth accounting is convenient because the growth rate in the TFP factor coincides with the trend growth rate of output per adult, namely the growth rate when hours worked h, the employment rate e, and the capital output ratio x (= K/Y) are constant.

Table 2.1 reports the growth rate of each of these factors for various subperiods since 1960. The capital share parameter θ is set at 0.362 (see our discussion below on calibration). The contribution of TFP growth between 1983–91 and 1991–2000 accounts for nearly all the decline in the growth in output per working-age person.[5] In spite of the low TFP growth in the 1973–83 period, output per adult increased at 2.2 percent. The reason why growth in output per adult was higher in the 1973–83 period than in the 1991–2000 period is that in the earlier period there was significantly more capital deepening and a smaller reduction in the labor input per working-age person.

Table 2.1 Accounting for Japanese growth per person aged 20–69 (%)

Period	Growth rate	Factors: TFP factor	Capital intensity	Workweek length	Employment rate
1960–1973	7.2	6.5	2.3	−0.8	−0.7
1973–1983	2.2	0.8	2.1	−0.4	−0.3
1983–1991	3.6	3.7	0.2	−0.5	0.1
1991–2000	0.5	0.3	1.4	−0.9	−0.4

Households

We model workweek length h as being exogenous prior to 1993 and endogenous thereafter. Following Hansen (1985) and Rogerson (1988), labor is indivisible so that a person either works h hours or does not work at all. There is a stand-in household with N_t working-age members at date t. The size of the household evolves over time exogenously. Measure E_t of the household members work a workweek of length h_t. The stand-in household utility function is:

$$\sum_{t=0}^{\infty} \beta^t N_t U(c_t,h_t,e_t) \quad \text{with} \quad U(c_t,h_t,e_t) = \log c_t = g(h_t)e_t, \tag{2.4}$$

where $e_t \equiv E_t/N_t$ is the fraction of household members that work and $c_t \equiv C_t/N_t$ is per member consumption.

As policy decreased the workweek length over time, the disutility of working depends on h. This disutility function is approximated in the neighborhood of $h = 40$ by a linear function:[6]

$$g(h) = \alpha[1 + (h - 40)/40]. \tag{2.5}$$

For this function, if not constrained, the workweek length chosen by the household is 40 hours. This follows from household first-order conditions (2.11) and (2.12), below.

To incorporate taxes, we assume that the only distorting tax is a proportional tax on capital income at rate τ. We could also incorporate a proportional tax on labor income. Provided that the rate is constant over time, the labor tax does not affect any of our results. This is because the labor tax, if included in the model, will be fully offset by a change in the calibrated value of α (see the consumption-leisure first-order condition (2.11) and (2.12), below, to see this point more clearly). Since 1984 there has been no major tax reform affecting income taxes, it is reasonable to assume that the average marginal tax rate on labor income (that is, the marginal tax rates averaged over different tax brackets) has been constant. We treat all other taxes as a lump-sum tax. The resulting period-budget constraint of the household, which owns the capital and rents it to the business sector, is:

$$C_t + X_t \leqslant w_t h_t E_t + r_t K_t - \tau(r_t - \delta)K_t - \pi_t. \tag{2.6}$$

Here w_t represents the real wage, π_t the lump-sum taxes and r_t is the rental rate of capital.

The after-tax interest rate equals:

$$i_t = (1 - \tau)\,(r_{t+1} - \delta). \tag{2.7}$$

The reason why we include a capital income tax is that a key variable in our analysis is the after-tax return on capital and this return is taxed at a high rate in Japan, even higher than in the United States.

Closing the model

Aggregate output Y_t is divided between consumption C_t, government purchases of goods and services G_t, and investment X_t.[7] Thus:

$$C_t + X_t + G_t = Y_t. \tag{2.8}$$

Capital depreciates geometrically, so:

$$K_{t+1} = (1 - \delta)\,K_t + X_t. \tag{2.9}$$

The government budget constraint is implied by the household budget constraint (2.6) and the resource constraint (2.8). By treating the capital tax income rate τ as a policy parameter, we are assuming that changes in government purchases are financed by changes in the lump-sum tax π_t. Thus, Ricardian equivalence holds in our model.

Calibration

We calibrate the model to the Japanese economy during 1984–89. There are five model parameters: θ (capital share in production), δ (depreciation rate), β (discounting factor), α (disutility of working) and τ (capital income tax rate). The data on the economy that go into the following calibration (such as data on taxes on capital income) are described in Appendix 2A1.

θ: The share parameter is determined in the usual way, as the sample average over the period 1984–89 of the capital income share in GNP.
δ: This is set equal to the sample average over the 1984–89 period of the ratio of depreciation to the beginning-of-the-year capital stock.
τ: This is set equal to the average rate in the 1984–89 period.
β: The discount factor is obtained from the intertemporal equilibrium condition:

$$\frac{U_{c_t}}{U_{c_{t+1}}} = \frac{c_{t+1}}{c_t} = \beta[1 + (1-\tau)(r_{t+1} - \delta)], \tag{2.10}$$

where U_{c_t} is the marginal utility of consumption for the period-utility function given in (2.4) and r_{t+1} is the marginal productivity of capital. We average this equation over the 1984–89 period and solve for β.
α: The disutility of work parameter α is obtained from the household maximization conditions for e and h:

$$c_t g(h_t) = w_t h_t \tag{2.11}$$

$$c_t g'(h_t) E_t = w_t E_t. \tag{2.12}$$

Equation (2.11) holds whether or not h is constrained and is the equation used to calibrate α. The calibrated value is the average value for the 1993–2000 period, the years that the workweek was not constrained.

The calibrated parameter values are displayed in Table 2.2.

Table 2.2 Calibration

Parameter	Value	Parameter	Value
θ	0.362	α	1.373
δ	0.089	τ	0.480
β	0.976		

Findings

We have calibrated the growth model to the Japanese economy for the 1984–89 period. We now use this calibrated model to predict what will happen in the 1990s and beyond.[8]

Initial conditions and exogenous variables

The simulation from year 1990 takes the actual capital stock in 1990 as the initial condition. The exogenous variables are (A_t, N_t, ψ_t), where ψ_t is G_t/Y_t, the GNP share of government purchases. We also take hours worked h_t to be exogenous for $t = 1990$–92. We need to specify the time path of those exogenous variables from 1990 on. For the 1990s ($t = 1990, 1991, \ldots, 2000$), we use their actual values. For $t = 2001, 2002, \ldots$, we assume the following. The TFP factor $A_t^{1/(1-\theta)}$ is set to its 1991–2000 average of 0.29 percent. We assume no population growth so that N_t is set to its 2000 value. The government's share ψ_t is set equal to its value in the 1999–2000 period of 15 percent.

Our simulation is deterministic. The issue of what TFP growth expectations to assign to the economic agents is problematic. We do not maintain that the decline in the growth rate of the TFP factor in the 1990s was forecast in 1990, even though we treat it as if it were. The justification is that a deterministic model is simple and suffices for answering our question of why the 1990s was a lost decade for the Japanese economy. If expectations had been modeled in any not unreasonable way, the key predictions of the model would be essentially the same. In particular, the magnitudes of the increase in the capital–output ratio and the fall in the return on capital would be the same.

Figures 2.6–2.8 report the behavior of the model and actual outcomes. As can be seen from Figure 2.6, the actual output in the 1990–2000 period is close to the predictions of our calibrated model. Theory with TFP exogenous predicts Japan's chronic slump in the 1990s.

The observed deepening of capital and the decline in the rate of return, noted in Section 2 and reproduced in Figures 2.7 and 2.8, are also predicted by the model. The capital–output ratio rises as output growth falls because the capital–output ratio associated with a lower productivity growth is higher. This can easily be seen from equation (2.10). In the new steady state with lower productivity growth, the consumption growth rate is lower, which means that the rate of return from capital is lower. Under diminishing returns to capital, the capital–output ratio must therefore be higher.

The difference in the precise paths of the model and actual path of the capital–output ratio is not bothersome given the model's assumption that the future path of the TFP factor was predicted perfectly by the economic

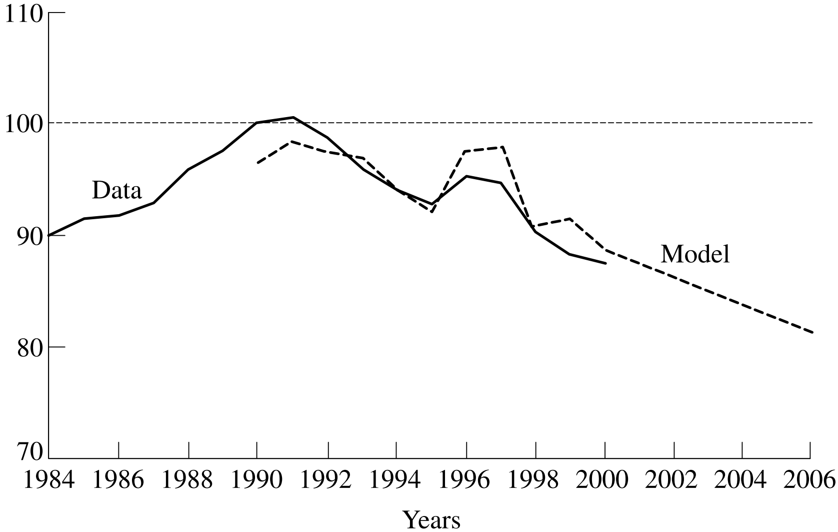


Figure 2.6 Detrended real GNP per working-age person (1990 = 100)

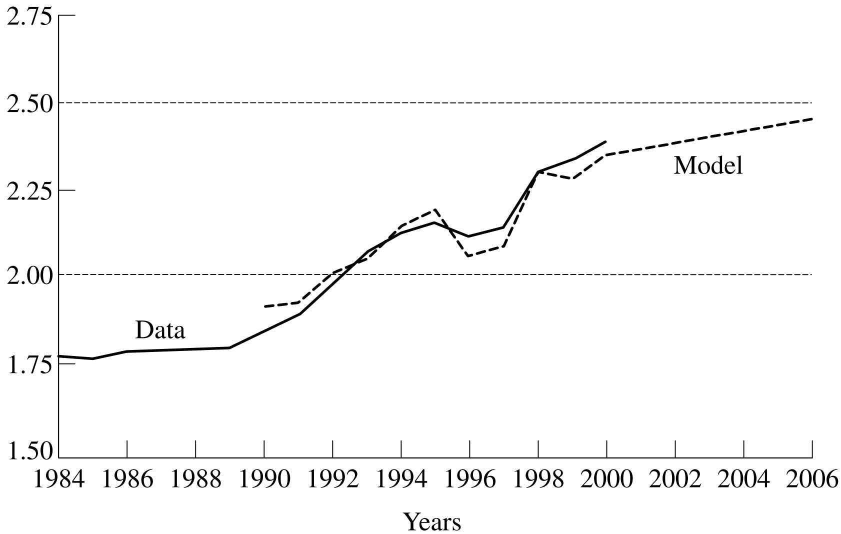


Figure 2.7 Capital–output ratio

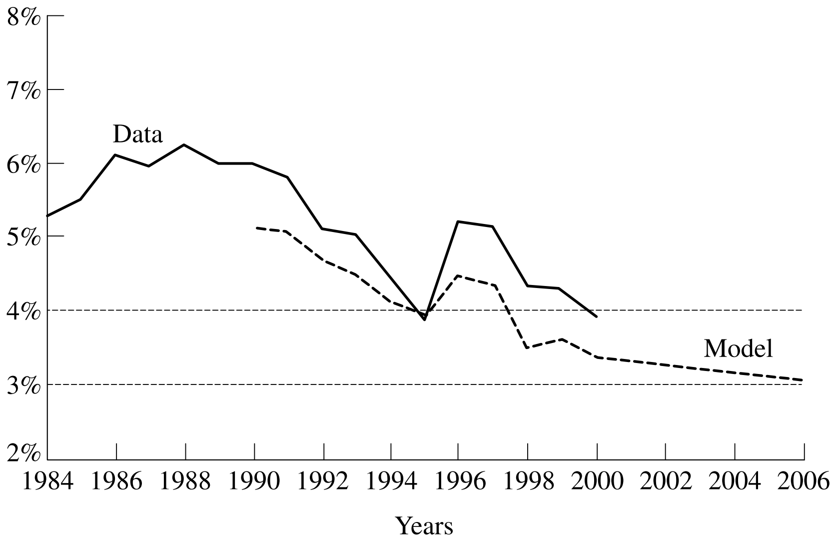


Figure 2.8 After-tax rate of return

agents when in fact it is not. Nor is the discrepancy between model and actual returns in Figure 2.8 bothersome. This is as expected given that actual returns include return on land as well as capital as discussed in Section 2.

The model's predictions for the 1990s are not sensitive to the values of the exogenous variables for the years beyond 2000. The predictions for the first decade of the twenty-first century, however, depend crucially on the values of the exogenous parameters for that decade. The most important variable is TFP. If the TFP growth rate increases to the historical norm of the industrial leader, Japan will not fall further behind the leader – rather, it will maintain its position relative to the industrial leader. If on the other hand, TFP growth is more rapid than the leader, Japan will catch up. We make no forecasts as to what TFP growth will be, and emphasize that this forecast is conditional on the TFP growth rate remaining low.

Assuming that TFP growth remains low, Japan cannot rely on capital deepening for growth in per-working-age-person output as it did in the past, as the Japanese capital stock is near its steady-state value. On the other hand, decreases in the labor input (aggregate hours) will not reduce growth as it has in the past, because, under our specification (2.5), average hours worked h will not magnify the disutility of aggregate hours worked when it is less than 40 hours. The Japanese people now work approximately the same number of hours as do Americans. If TFP growth again becomes as rapid as it was in the 1983–91 period, the labor input will increase and this will have a positive steady-state level effect on output.

4. WAS INVESTMENT CONSTRAINED?

An important alternative hypothesis about Japan's lost decade is what we call the 'credit crunch' hypothesis. It holds that, for one reason or another, there is a limit on the amount a firm can borrow. If bank loans and other means of investment finance are not perfect substitutes, an exogenous decrease in the loan limit constrains investment and hence depresses output.[9] This hypothesis is becoming an accepted view even among academics. It has an appeal because the collapse of bank loans and the output slump occurred in the same period (the 1990s) and because the collapse of bank loans seems exogenous, taking place when the BIS capital ratio is said to be binding for many Japanese banks. In this section, we confront this 'credit crunch' hypothesis with data from various sources.

Evidence from the National Accounts

As mentioned at the end of Section 2, the output share of domestic investment declined substantially in the 1990s. If this decline is due to reduced bank lending, we should see much of the decline in investment by nonfinancial corporations. The Japanese National Accounts has a flow-of-funds account (called the capital transactions account) for the nonfinancial corporate sector that allows us to examine sources of investment finance. The cash flow identity for firms states that:

investment (excluding inventory investment) (2.13)
= (a) net increase in bank loans
+ (b) net sales of land
+ (c) gross corporate saving (that is, retention plus accounting depreciation)[10]
+ (d) net increase in other liabilities (that is, new issues in shares and corporate bonds plus net decrease in financial assets).

The capital transactions account in the National Accounts allows one to calculate items (a)–(d) above for the nonfinancial corporate sector.[11]

Figure 2.9 shows investment (excluding inventory investment) and item (a) (net increase in bank loan balances) as ratios to GNP. (The difference between the two, of course, is the sum of items (b), (c) and (d).) There are two things to observe. First, the dive in the output share of domestic investment, shown in Figure 2.5, did not occur in the nonfinancial corporate sector. The output share of investment by nonfinancial corporations

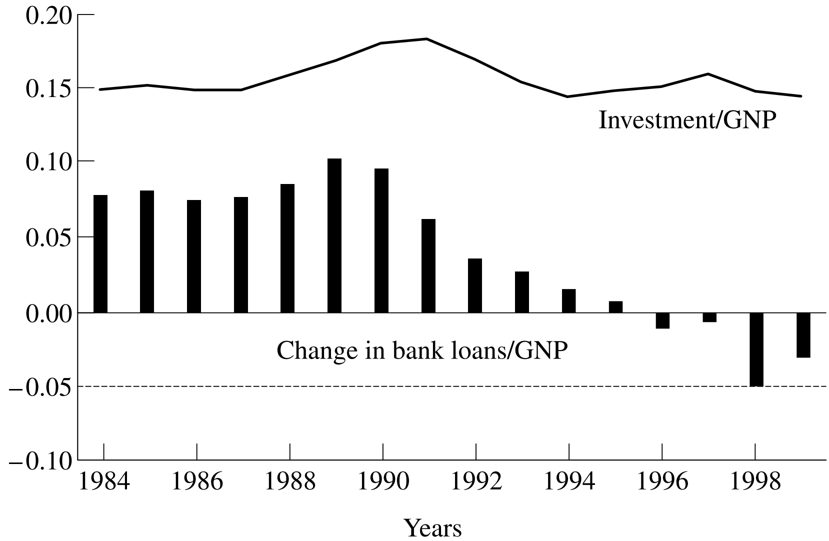


Figure 2.9 Collapse of bank loans: nonfinancial corporate sector

Table 2.3 Sources of investment finance for nonfinancial corporations (%)

Sources of fund as fraction of investment	1984–1988	1993–1999
(a) Bank loans	52.2	−4.8
(b) Sale of land	−6.9	5.7
(c) Gross savings	79.2	88.1
(d) Net increase in other liabilities	−24.5	11.0
Total	100	100

remained at 15 percent, except for the 'bubble' period of the late 1980s and early 1990s when the share was higher. Second, investment held up despite the collapse of bank loans in the 1990s.[12] That is, other sources of funds replaced bank loans to finance the robust investment by nonfinancial corporations in the 1990s. To corroborate this second point, Table 2.3 shows how the sources of investment finance changed from 1984–88 to 1993–99 (thus excluding the bubble period). In the 1980s, bank loans and gross corporate saving financed not only investment but also purchases of land (see the negative entry for 'sale of land' in the table) and a buildup of financial assets (see the negative entry for 'net increase in other liabilities'). In the 1990s, firms drew down the land and financial assets that had been built up during the 1980s to support investment. These observations are inconsistent with the credit crunch hypothesis.

Evidence from Survey Data on Private Nonfinancial Corporations

The preceding discussion, based on the National Accounts data, ignores distributional aspects. For example, large firms may not have been constrained while small ones were. As is well known (see, for example, Hoshi and Kashyap 1999), as a result of the liberalization of capital markets, large Japanese firms scaled back their bank borrowing and started to rely more heavily on open-market funding, and the shift away from bank loans is complete by 1990. It is also well known that for small firms, essentially the only source of external funding is still bank loans. Therefore, if investment is constrained for some firms, those firms must be small firms. How did the collapse of bank loans affect small firms?

The most comprehensive survey of private nonfinancial corporations (a subset of the nonfinancial corporate sector examined above) in Japan is a survey by the Ministry of Finance (MOF).[13] From annual reports of this survey published by the MOF, sample averages of various income and balance sheet variables for 'small' firms (whose paid-in capital is less than 1 billion yen) can be obtained for fiscal years (a Japanese fiscal year is from April of the year to March of the next year). Figure 2.10 is the small-firm version of Figure 2.9. The difference between investment and bank loans in the 1980s is much smaller in Figure 2.10 than in Figure 2.9, underscoring the importance of bank loans for small firms. In the 1990s, however, as in Figure 2.9, investment held firm in spite of the collapse of bank loans. The sources of investment finance for small firms are shown in Table 2.4. It is not meaningful in the MOF survey to distinguish between items (c) (gross saving) and (d) (net increase in liabilities other than bank loans) in the cash-flow identity (2.13). For example, suppose the firm reports hitherto unrealized capital gains on financial asset holdings by selling those assets and then immediately buying them back. This operation increases (c) and decreases (d) by the same amount. Therefore, in Table 2.4, items (c) and (d) are bundled into a single item called 'other.' The table shows that small firms, despite the collapse of bank loans, continued to increase land holdings in the 1990s. That is, gross corporate saving and net decreases in financial assets combined were enough to finance not only the robust level of investment but also land purchases – as all the while the loan balance was being reduced.

As just noted, it is not possible to tell from the MOF survey which component – saving or a running down of assets – contributed more. It is, however, instructive to examine the evolution of a component of financial

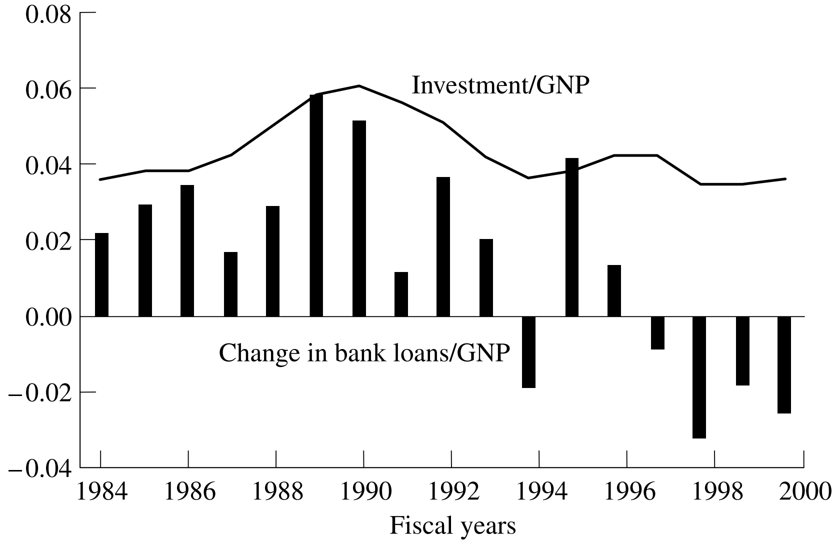


Figure 2.10 Investment finance: small firms

Table 2.4 Sources of investment finance for small nonfinancial corporations (%)

Sources of fund as fraction of investment	Fiscal year 1984–88	Fiscal year 1993–2000
(a) Bank loans	64.5	−12.6
(b) Sale of land	−18.3	−20.8
(c) + (d) Other	53.8	133.4
Total	100	100

Note: *Rounded.

assets whose reported value cannot be distorted by inclusion of unrealized capital gains. Figure 2.11 graphs the ratio of cash and deposits to the (book value of) capital stock. First of all, the ratio is huge. The ratio for the nonfinancial corporate sector as a whole in the National Accounts is about 0.4. In contrast, the US ratio for nonfinancial corporations is much lower, less than 0.2, according to the Flow of Funds Accounts compiled by the Board of Governors. For some reason the ratio was high in the early 1980s.[14] It is clear from this and the previous figure that small firms during the bubble period used the cash and bank loans for financial investments.

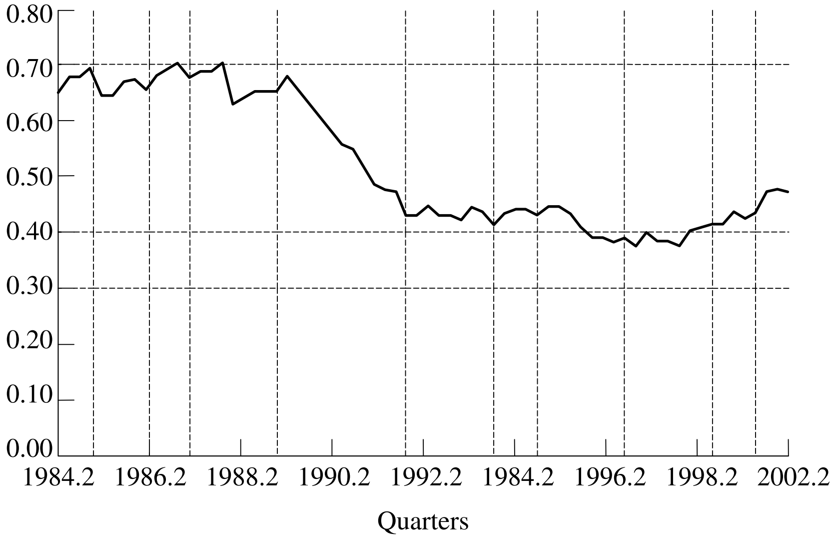


Figure 2.11 Ratio to cash plus deposits to capital stock: small firms

Second, turning to the mid- to late 1990s, Figure 2.10, above, indicates that small firms relied on cash and deposits as a buffer against the steep decline in bank loans.

Evidence from Cross-section Regressions

In the early 1990s, there was an active debate in the United States about whether the recession in that period was due to a credit crunch. To answer this question, Bernanke and Lown (1991) examined evidence from US states on output and loan growth. Based on a variety of evidence, including a cross-section regression involving output and loan growth by state, they concluded that the answer is probably 'no'. In this subsection, we estimate the same type of regression for the 47 Japanese prefectures.

For the recession period of 1990–91, Bernanke and Lown find that employment growth in each state is related to contemporaneous growth in bank loans, with the bank loan regression coefficient of 0.207 with a *t* value of over 3. A positive coefficient in the regression admits two interpretations. The first is the credit crunch hypothesis that an exogenous decline in loan supply constrains investment and hence output. The second is that the observed decline in bank loans is due to a shift in loan demand. Bernanke and Lown prefer the second interpretation because the positive coefficient became insignificant when loan growth is instrumented by the capital ratio.

For Japan, we have available GDP by prefecture for fiscal years (April to March of the following year) and loan balance (to all firms and also to small firms whose paid-in capital is 100 million yen or less) at the end of March of each year.[15] The regression we run across prefectures is:

$$\text{GDP growth rate} = \beta_0 + \beta_1 \cdot \text{bank loan growth rate.} \qquad (2.14)$$

According to the official dating of business cycles (published by the ESRI: Economic and Social Research Institute of the Cabinet Office of the Japanese government), there have been five recessions since 1975: from March 1977 to October 1977, from February 1980 to February 1983, from June 1985 to November 1986, from February 1991 to October 1993, and from March 1997 to April 1999. Without monthly data, it is not possible to align these dates with our data on GDP and bank loans. We therefore focus on the three longer recessions.

Our results are reported in Table 2.5. In the regression for 1996–98, for example, the dependent variable is GDP growth from fiscal year 1996 (April 1996 to March 1997) to fiscal year 1998 (April 1998 to March 1999). This GDP growth is paired with the growth in loan balance from March 1996 to March 1998.[16] The loan growth is for all firms in Regression 1 and for small firms in Regression 2. Regression 1 is comparable to the state-level regression in Bernanke and Lown (1991) for the US states, except that the measure of output growth here is GDP growth, not employment growth.[17] Overall, the loan growth coefficient is not significant, which is consistent with our view that there may have been a credit crunch but it did not matter for investment because firms found other ways to finance investment.

Table 2.5 Cross-section regression of GDP growth on loan growth

Recession years	Regression 1	Regression 2
	Independent variable is loan growth to all firms from March to March over indicated years	Independent variable is loan growth to small firms from March to March over indicated years
1979–1982	0.046 (0.3)	0.125 (0.9)
1990–1993	0.090 (1.0)	0.049 (0.6)
1996–1998	0.125 (2.0)	0.120 (1.7)

Note: *t*-values in parentheses. The dependent variable is GDP growth rate over indicated fiscal years. The coefficient on the constant in the regression is not reported.

The significant coefficient for 1996–98 suggests that the recession in the late 1990s was partly due to a credit crunch, but this period is special. The three-month commercial paper rate, which has been about 0.5 to 0.6 percent since January 1996, shot up to above 1 percent in December 1997 and stayed near or above 1 percent before coming down to the 0.5 to 0.6 percent range in April 1998. During this brief period, various surveys of firms (for example, the Bank of Japan's survey, *Tankan Survey*) report a sharp rise in the fraction of small firms that said it was difficult to borrow from banks. The regression result in Table 2.5, which detects a significant association between output and bank loans for 1996–98 but not for other periods, gives us confidence that the credit crunch hypothesis, while possibly relevant for output for a few months from late 1997 to early 1998, cannot account for the decade-long stagnation.[18]

5. CONCLUDING COMMENTS

In examining the virtual stagnation that Japan began experiencing in the early 1990s, we find that the problem is not a breakdown of the financial system, as corporations large and small were able to find financing for investments. There is no evidence of profitable investment opportunities not being exploited due to lack of access to capital markets. Those projects that are funded are on average receiving a low rate of return.

The problem is low productivity growth. If it remains lower in Japan than in the other advanced industrial countries, Japan will fall further behind. We are not predicting that this will happen and would not be surprised if Japanese productivity growth returned to its level in the 1984–89 period. We do think that research effort should be focused on determining what policy reform will allow productivity to again grow rapidly.

We can only conjecture on what reforms are needed. Perhaps the low productivity growth is the result of a policy that subsidizes inefficient firms and declining industries. This policy results in lower productivity because the inefficient producers produce a greater share of the output. This also discourages investments that increase productivity. Some empirical support for this subsidizing hypothesis is provided by the experience of the Japanese economy in the 1978–83 period. During that five-year period when the 1978 'Temporary Measures for Stabilization of Specific Depressed Industries' law was in effect (see Peck et al. 1988), the TFP growth rate was a dismal 0.64 percent. In the three years prior, the TFP growth averaged 2.18 percent and in the six-year period after, it averaged slightly over 2.5 percent.

We said very little about the bubble period of the late 1980s and early 1990s, a boom period when property prices soared, investment as a fraction of GDP was unusually high, and output grew faster than in any other years in the 1980s and 1990s. We think the unusual pickup in economic activities, particularly investment, was due to an anticipation of higher productivity growth that never materialized. To account for the bubble period along these lines, we need to have a model where productivity is stochastic and where agents receive an indicator of future productivity. But the account of the lost decade by such a model would be essentially the same as the deterministic model used in this chapter.

APPENDIX 2A1 DATA

This appendix is divided into two sections. In the first section, we describe in detail how we constructed the model variables used in our neoclassical growth model. The second section describes how the data underlying the tables and figures in the text are constructed. All the data are in Excel files downloadable from www.e.u-tokyo.ac. jp/~hayashi/hp.

Construction of Model Variables

The construction can be divided into two steps. The first is to make adjustments to the data from the Japanese National Accounts, which is our primary data source, to make them consistent with our theory. The second step is to calculate model variables from the adjusted national accounts data and other sources. The exact formulas of these steps can be found in the Excel file 'rbc.xls' downloadable from the URL mentioned above.

Step 1: adjustment to the National Accounts

Various adjustments to the National Accounts are needed for three reasons. First, depreciation in the National Accounts is on historical cost basis. Second, in our theory all government purchases are expensed. Third, starting in 2001 the National Accounts (compiled by the ESRI) adopted a new standard (called the 1993 SNA (System of National Accounts) standard) that is different from the previous standard (the 1968 SNA).

Extension to 1999 and 2000 For years up to 1998, the *2000 Annual Report on National Accounts* has consistent series under the 1968 SNA standard. The *2001 Annual Report*, which adopted the 1993 SNA standard, has series only for 1991–99. The ESRI also releases series on the 1968 SNA basis for years up to 2000, but those series are only for a subset of the variables forming the income and product accounts. Furthermore, those accounts divide the whole economy into subsectors in a way different from the sector division in the *2000 Annual Report*. From these three sources, it is possible, under the usual sort of interpolation and extrapolation, to construct consistent series for all relevant variables under the 1968 SNA standard up to 2000 (consult the Excel file mentioned above for more details). On the left-hand side of Table 2A1.1, we report values (relative to GNP) of items in the income and product accounts thus extended to 2000, averaged over 1984–2000. Also reported are capital stocks relative to GNP. Beginning-of-year (end-of-previous-year) capital stocks for years up to 1999 are directly available from the *2000 Annual Report*; capital stocks at the beginning of 2000 are taken from the *2001 Annual Report*.

Capital consumption adjustments The Japanese National Accounts include the balance sheets as well as the income and product accounts for the subsectors of the economy. In the income and product accounts, depreciation (capital consumption) is on historical cost basis, while in the balance sheets, capital stocks are valued at replacement costs. As was pointed out in Chapter 11 of Hayashi (1997), replacement cost depreciation implicit in the balance sheets can be estimated – under a certain set of assumptions – from various accounts included in the National Accounts. For years up to 1998, this Hayashi estimation of replacement cost depreciation is possible from the *2000 Annual Report*, which conforms to the 1968 SNA standard and which includes data for years up to 1998. The procedure is in the Excel file 'japsave.xls', downloadable from the URL mentioned above. The *2001 Annual Report*, which adopted the 1993 SNA standard, actually reports replacement cost depreciation in its balance sheet section for 1991–99. However, since the class of assets in the new SNA is broader, we use only the 1999 value and use it only to obtain our estimate of the 1999 value from the 1998 Hayashi estimate. For 2000, we linearly extrapolate from the 1998 and 1999 numbers. Consult the Excel file 'rbc.xls' mentioned above for more details. From the estimate of replacement cost depreciation, an estimate of capital consumption adjustment can be obtained as the difference between the replacement cost depreciation thus calculated and the historical cost depreciation reported in the National Accounts. We use this capital consumption adjustment to make the National Account variables consistent with replacement cost accounting. For example, we add this capital consumption adjustment to (book value) depreciation to obtain depreciation at replacement costs, and we subtract the capital consumption adjustment from operating surplus.

Treatment of government capital In our theory, all government purchases are expensed. Consequently, government consumption in the product account includes government investment, and capital consumption on government capital is subtracted from GNP to define (adjusted) GNP.

These two adjustments, capital consumption adjustments and expensing of government investment, are shown on the right-hand side of Table 2A1.1, where we provide descriptions of the adjustments and the adjusted values (relative to the unadjusted GNP).

Step 2: calculation of model variables from the adjusted National Accounts
The variables of our model are the following: W (wage income), R (capital income), DEP (depreciation), Y (adjusted GNP, exclusive of capital consumption on government capital), C (private consumption), X (investment, domestic investment plus investment in foreign assets), G (government consumption), K (capital stock), h (hours worked per employed person), E (number of employed persons), N (working age population), and taxes on capital income. Of these, W, R and DEP are used to evaluate the capital income share θ as described in Section 3 ('Calibration') of the text.

Income and product account variables Table 2A1.2 explains how the variables comprising the income and product accounts are constructed exclusively from the adjusted National Accounts. Imputed rent, which is the housing component of operating surplus in the noncorporate sector, is included in capital income. We assume that 80 percent of operating surplus in the nonhousing component of the noncorporate sector is wages. We need to divide indirect taxes between wages and capital income. For lack of good alternatives, we simply split it in half. Statistical discrepancy is allocated proportionately between W, R and DEP. Thus, by construction, the sum of W, R and DEP equals Y (GNP exclusive of capital consumption on government capital).

Capital stock, K Capital stock excludes government capital but includes capital in foreign countries. Capital in Foreign Countries (KF) was calculated in the following way: $KF(1989) = 25 *$ Net Factor Payments(1989), $KF(t+1) = KF(t) +$ Net Exports(t) + Net Factor Payments(t).

Average hours worked, h This variable is from an establishment survey conducted by the Ministry of Welfare and Labor (this survey is called *Maitsuki Kinro Tokei Chosa*). We use a series, included in this survey, for establishments with 30 or more employees. (There is a series for establishments with five or more employees, but it has been available only since 1990.)

Employment, E The number of employed persons for 1970–98 is available from the National Accounts (see Table I-[3]-3 of the *2000 Annual Report on National Accounts*). The *Labor Force Survey* (compiled by the General

Table 2A1.1 National Income Accounts adjustments

National income accounting concept	Value	Adjustments	Value
Income			
1. Compensation of employees	0.546		0.546
2. Operating surplus	0.223		0.198
Corporate	0.105	− Adjustment of capital consumption in corporate (0.014)	0.091
Noncorporate	0.118		0.107
– Housing	0.050	− 70% adjustment of capital consumption in noncorporate (0.008)	0.042
– Nonhousing	0.068	− 80% adjustment of capital consumption in noncorporate (0.003)	0.065
3. Capital consumption	0.151		0.170
Government	0.006	− Capital consumption in government (0.006)	0.000
Corporate	0.099	+ Adjustment of capital consumption in corporate (0.014)	0.114
Noncorporate	0.045	+ Adjustment of capital consumption in noncorporate (0.011)	0.056
4. Indirect business taxes	0.072		0.072
5. Net factor payments	0.008		0.008
6. Statistical discrepancy	0.000		0.000
Total income	1.000		0.994
Product			
1. Consumption	0.684		0.732
Private	0.589		0.589
Government	0.095	+ Fixed capital formation in gov't – capital consumption in gov't (0.049)	0.143

2. Investment	0.288		0.233
Inventory	0.004		0.004
Fixed capital	0.285		0.230
– Government	0.055	– Fixed capital formation in government (0.055)	0.000
– Private (corporate plus noncorporate)	0.230		0.230
3. Current account	0.028		0.028
Net exports	0.020		0.020
Net factor payments	0.008		0.008
Total product	1.000		0.994
Capital stocks			
1. Government	0.647	– Government fixed assets (0.647)	0.000
2. Corporate	1.031		1.031
3. Noncorporate	0.575		0.575
4. Inventories, corporate	0.148		0.148
5. Inventories, noncorporate	0.021		0.021
6. Foreign capital	0.000	+ Net capital stock abroad (0.221)	0.221
Total capital stock	2.422		1.996

Note: Averages of ratios to unadjusted GNP over 1984–2000.

Table 2A1.2 Model variables and relation to adjusted NIA data

Variable	Name	Components
W	Wage income	Compensation of Employees + 0.8* Operating Surplus in Nonhousing Noncorporate Sector + 0.5* Indirect Business Taxes + Proportion of Statistical Discrepancy
R	Capital income	Operating Surplus in Corporate Sector + Operating Surplus in Housing Noncorporate Sector + 0.2* Operating Surplus in Nonhousing Noncorporate Sector + 0.5* Indirect Business Taxes + Proportion of Statistical Discrepancy + Net Factor Payments
DEP	Depreciation	Total Capital Consumption (Corporate + Noncorporate) + Proportion of Statistical Discrepancy
Y	Income = output	$W+R+DEP=Y=C+G+X$
C	Private consumption	Private consumption
G	Gov't expenditure	Adjusted Government Consumption
X	Investment	Total Investment (Corporate + Noncorporate) + Net Exports + Net Factor Payments
K	Capital stock	Total Capital Stock (Corporate + Noncorporate + Stock of inventories) + Capital in Foreign Countries

Affairs Agency) provides a different estimate of employment from 1960 to the present. To extend the estimate in the NIA back to 1960, we multiply the *Labor Force Survey* series by the ratio of the National Accounts estimate to the *Labor Force Survey* estimate for 1970.

Working-age population, N The working-age population is defined as the number of people between ages 20 and 69.

Taxes on capital income This variable is used to calculate the tax rate on capital income, denoted τ in the text. It is defined as the sum of direct taxes on corporate income (available from the income account for the corporate sector in the National Accounts), 50 percent of indirect business taxes, and 8 percent of operating surplus in the nonhousing component of the noncorporate sector.

Data Underlying Tables and Figures

Figures 2.1–2.5 and Table 2.1 use the model variables described in the first section of this appendix. Figures 2.6–2.8 are based on the simulation described in Section 3 of the text. The underlying data are in Excel file 'rbc.xls'.

- *Figure 2.9* Data on investment and bank loans are from the capital transactions account for nonfinancial corporations in the National Accounts (Table 1-[2]-III-1). For 1984–98, the data are from the *2000 Report on National Accounts*, and the GNP used to deflate investment and bank loans are constructed as in the first section of this appendix. For 1999, the data are from the *2001 Report on National Accounts*. The GNP for 1999 used to deflate is directly from this report. This is because the definition of investment in the 2001 report is based on the 1993 SNA definition. The data underlying this figure and Table 2.3 are in Excel file 'nonfinancial.xls' downloadable from the URL already mentioned.
- *Table 2.3* This too is calculated from the capital transactions account for nonfinancial corporations, available from the *2000 Report* (for data for 1984–98) and the *2001 Report* (for 1999). Investment (excluding inventory investment), gross saving (defined as net saving plus depreciation), bank loans and sale of land are directly available from the capital transactions account. Net increase in other liabilities is defined as investment less the sum of bank loans, sale of land and gross saving. So the net increase in other liabilities, bank loans, sale of land and gross saving add up to investment.
- *Figure 2.10* The data source is *Hojin Kigyo Tokei* (*Incorporated Enterprise Statistics*) collected by the MOF. It is a large sample (about 18 thousand) of corporations from the population of about 1.2 million (as of the first quarter of 2000) listed and unlisted corporations excluding only very tiny firms (those with less than 10 million yen in paid-in capital). In the second quarter of each year, a freshly drawn sample of firms report quarterly income and balance-sheet items for four consecutive quarters comprising the fiscal year (from the second quarter of the year to the first quarter of the next year). The sampling ratio depends on firm sizes, with a 100 percent sampling of all 'large' firms (about 5400 firms, as of fiscal year 2000) whose paid-in capital is 1 billion yen or more. The MOF publishes sample averages by firm size. The sample averages we use are for 'small' firms whose paid-in capital is less than 1 billion yen. For each fiscal year (April of the calendar year to March of the following

year), investment for the fiscal year is the sum over the four quarters of the fiscal year of the sample average of investment (excluding inventory investment). The net increase in bank loans for fiscal year t is the difference in the loan balance (defined as the sum of short-term and long-term borrowings from financial institutions) between the end of fiscal year t (that is, the end of the first quarter of calendar year $t + 1$) and the end of the previous fiscal year (that is, the end of the first quarter of calendar year t). Information on the balance sheet at the end of the previous fiscal year is available because the MOF collects this information for the firms newly sampled in the second quarter of year t. The GNP used to deflate is constructed as described in the first section of this appendix. The data underlying this figure, Table 2.4, and Figure 2.11 are in Excel file 'mof.xls' downloadable from the URL already mentioned.

- *Table 2.4* The MOF survey is the source of this table also. Calculation of investment and bank loans is already described above for Figure 2.10. Sale of land for fiscal year t is the difference in the book value of land between the end of fiscal year t and the end of the previous fiscal year. The value for 'other' is calculated as investment less the sum of bank loans and sale of land.
- *Figure 2.11* This too is calculated from the MOF survey. It is the ratio of the sample average of cash and deposits for the small firms to the corresponding sample average of the book value of fixed assets (excluding land) at the end of each quarter.
- *Table 2.5* Data on prefectural GDP for fiscal years are available from the *Report on Prefectural Accounts* (various years) published by the ESRI. Loan balance for domestically chartered banks by prefecture at the end of each March is available from *A Survey on Domestically Chartered Bank Lending by Prefecture and by Client Firm's Industry* by the Statistics Department of the Bank of Japan. The underlying data are in 'prefecture.xls' downloadable from the URL already mentioned.

APPENDIX 2A2 COMPUTATIONAL ISSUES

This appendix describes the solution procedure in detail and provides additional findings.

Equilibrium Conditions

In the text, we wrote the disutility of work function as $g(h_t)$ (where h_t is the workweek length or hours worked per week). To make explicit that it also depends on some exogenous variable z_t, we will write the function as:

$$\text{disutility of work: } g(h_t;\, z_t). \tag{2A2.1}$$

In our analysis below, we take z_t to be the number of days worked per week. Government actions since 1988 (described in the text), which increased the number of national holidays by three and which made it costly for people to work on Saturdays by requiring banks and government offices to be closed on Saturdays, are a good reason to assume that the number of days worked per week is an exogenous variable determined outside the economic model.

Otherwise the notation in this appendix is the same as in the text: Y_t = aggregate output in period t, A_t = the level of TFP, K_t = capital stock, h_t = hours worked, N_t = working-age population, E_t = employment, e_t = employment rate defined as $e_t \equiv E_t/N_t$, r_t = rental rate of capital, w_t = hourly real wage rate, C_t = aggregate consumption, X_t = investment, G_t = government expenditure, τ = tax rate on capital income, and (as just mentioned) z_t = days worked per week.

To reproduce relevant equations from the text,

$$\text{aggregate production function: } Y_t = A_t K_t^{\theta}(h_t e_t N_t)^{1-\theta}, \tag{2A2.2}$$

$$\text{marginal productivity condition for capital: } r_t = \theta A_t K_t^{\theta-1}(h_t e_t N_t)^{1-\theta}, \tag{2A2.3}$$

$$\text{marginal productivity condition for labor: } w_t = (1-\theta) A_t K_t^{\theta}(h_t e_t N_t)^{-\theta}, \tag{2A2.4}$$

$$\text{resource constraint: } C_t + X_t + G_t = Y_t, \tag{2A2.5}$$

$$\text{capital accumulation: } K_{t+1} = (1-\delta)K_t + X_t, \tag{2A2.6}$$

$$\text{Euler equation:} \quad \frac{C_{t+1}}{N_{t+1}} = \frac{C_t}{N_t} \cdot \beta \cdot [1 + (1-\tau)(r_{t+1} - \delta)], \quad (2A2.7)$$

$$\text{first-order condition for } e_t: \quad \frac{C_t}{N_t} g(h_t; z_t) = w_t h_t, \quad (2A2.8)$$

$$\text{first-order condition for } h_t: \quad \frac{C_t}{N_t} g'(h_t; z_t) = w_t. \quad (2A2.9)$$

Here, $g'(h_t; z_t)$ is the derivative with respect to h_t.

Substituting the production function (2A2.2), the marginal productivity conditions ((2A2.3) and (2A2.4)), and the capital accumulation equation (2A2.6) into the household's first-order conditions ((2A2.7) and (2A2.8)) and the resource constraint (2A2.5), we obtain the following equilibrium conditions:

$$\frac{C_t}{N_t} g(h_t; z_t) = (1-\theta) A_t K_t^{\theta} (h_t e_t N_t)^{-\theta} h_t, \quad (2A2.10)$$

$$\frac{C_{t+1}}{N_{t+1}} = \frac{C_t}{N_t} \cdot \beta \cdot \{1 + (1-\tau)[\theta A_{t+1} K_{t+1}^{\theta-1} (h_t e_{t+1} N_{t+1})^{1-\theta} - \delta]\}, \quad (2A2.11)$$

$$K_{t+1} = (1-\delta) K_t + A_t K_t^{\theta} (h_t e_t N_t)^{1-\theta} - C_t - G_t. \quad (2A2.12)$$

These equations together determine the sequence of $\{C_t, K_t, e_t\}$ given the sequence of $\{A_t, N_t, G_t, h_t, z_t\}$ as well as the initial value for K_t.

Derivation of the Disutility of Work Function

Combining (2A2.8) and (2A2.9), we obtain:

$$h_t g'(h_t; z_t) = g(h_t; z_t). \quad (2A2.13)$$

As explained in the next paragraphs, $g(h_t; z)$ is convex in h and $g(0, z) > 0$ because of the fixed cost of going to work, which means that condition (2A2.13) uniquely determines h_t. This is described in Figure 2A2.1, where at point A the slope g' equals the ratio g/h. Write this unique solution as $h(z_t)$. Thus, in our model, hours worked are determined separately, independent of the TFP level A_t.

To see why g is convex and $g(0; z_t) > 0$, we recall from the literature on indivisible labor Hansen 1985 and Rogerson 1988 how the function is related to the utility of leisure. Let $\phi(h_t; z_t)$ be the utility of leisure as a function of hours worked per week h_t. It is decreasing and concave in h_t, so $\phi'(h;$

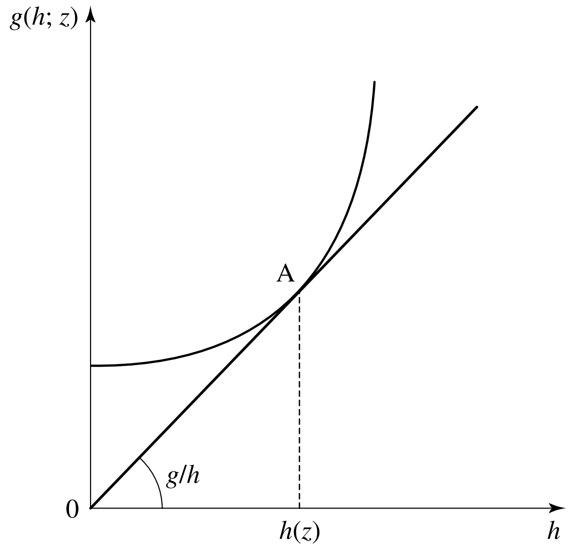


Figure 2A2.1 The work disutility function

$z) < 0$ and $\phi''(h; z) < 0$. Days worked, z_t, enter the utility function ϕ for two reasons. First, the fixed cost of working, which reduces the utility of leisure, should depend on z_t. Second, the number of days to work per week affects how hours worked are allocated over the week (see the example below for an illustration of this point).

In the indivisible labor model, the employment rate e_t is also the probability that the stand-in household gets employed. Thus the period utility function is:

$$\log(c_t) + \phi(h_t; z_t)e_t + \phi_0 \cdot (1 - e_t), \quad (2A2.14)$$

where ϕ_0 is the utility of leisure when the person does not work at all. This ϕ_0 should be greater than $\phi(0; z_t)$ for $z_t > 0$, because the latter is the utility from leisure if the person incurs a fixed cost of going to work by showing up at the workplace, only to work zero hours. The period utility (2A2.14) can be rewritten as (ignoring a constant term):

$$U(c_t, h_t, e_t; z_t) \equiv \log(c_t) - g(h_t; z_t)e_t, \text{ where } g(h_t; z_t) \equiv \phi_0 - \phi(h_t; z_t). \quad (2A2.15)$$

Since $\phi(h_t; z_t)$ is concave, $g(h_t; z_t)$ is convex in h_t. By construction, we have $g(0; z_t) = \phi_0 - \phi(0; z_t) > 0$.

As an example, consider a function of the form:

$$\phi(h_t; z_t) = \mu\left(1 - \frac{h_t}{z_t}\right) z_t + \mu(1)(1 - z_t) - \nu z_t \text{ and } \phi_0 = \mu(1). \tag{2A2.16}$$

Here, μ is the utility of leisure per day (with daily time endowment normalized to unity) as a function of daily leisure hours and $\nu(>0)$ is the fixed cost of commuting per working day. The g function is:

$$g(h_t; z_t) = \left[\nu + \mu(1) - \mu\left(1 - \frac{h_t}{z_t}\right)\right] z_t. \tag{2A2.17}$$

We have $g(0; z_t) = -\nu z_t > 0$ for $z_t > 0$ as required. It is easy to see that $h(z)$, determined by (2A2.13), is proportional to z (so hours worked per day, $h(z)/z$, does not depend on z).

Detrending

We can detrend the model by defining:

$$\tilde{k}_t \equiv \frac{K_t}{A_t^{\frac{1}{1-\theta}} N_t},\ \tilde{c}_t \equiv \frac{C_t}{A_t^{\frac{1}{1-\theta}} N_t},\ \tilde{y}_t \equiv \frac{Y_t}{A_t^{\frac{1}{1-\theta}} N_t},\ \tilde{\gamma}_t \equiv \frac{A_{t+1}^{\frac{1}{1-\theta}}}{A_t^{\frac{1}{1-\theta}}},\ \psi_t \equiv \frac{G_t}{Y_t},\ n_t \equiv \frac{N_{t+1}}{N_t}. \tag{2A2.18}$$

Also define the detrended capital–labor ratio (K/hE deflated by the TFP factor $A_t^{1/(1-0)}$) as:

$$\text{detrended capital–labor ratio: } x_t \equiv \frac{\tilde{k}_t}{e_t h(z_t)}. \tag{2A2.19}$$

Then, noting that $h_t = h(z_t)$, the above three equilibrium conditions (2A2.10)–(2A2.12) become:

$$\tilde{c}_t g[h(z_t); z_t] = (1-\theta) h(z_t) x_t^{\theta} \text{ or } x_t = \left\{\frac{\tilde{c}_t}{(1-\theta)\,\frac{h(z_t)}{g[h(z_t);z_t]}}\right\}^{1/\theta}, \tag{2A2.20}$$

$$\tilde{c}_{t+1} = \frac{\tilde{c}_t}{\gamma_t} \beta \cdot [1 + (1-\tau)\,(\theta x_{t+1}^{\theta-1} - \delta)], \tag{2A2.21}$$

$$\tilde{k}_{t+1} = \frac{1}{\gamma_t n_t}\,[(1-\delta) + (1-\psi_t) x_t^{\theta-1}]\tilde{k}_t - \tilde{c}_t. \tag{2A2.22}$$

These three equations determine $\{x_t, \tilde{c}_t, \tilde{k}_t\}$ given $\{\gamma_t, n_t, \psi_t, z_t\}$ and given the function $g(h; z)$ (which generates $h_t = h(z_t)$ via (2A2.13)). The employment rate e_t can be recovered from the definition of x_t by the formula:

$$e_t = \frac{\tilde{k}_t}{h(z_t) x_t}. \tag{2A2.23}$$

Detrended output $\tilde{y}_t$ can be calculated as:

$$\tilde{y}_t = \tilde{k}_t^{\theta}[h(z_t)e_t]^{1-\theta} = \tilde{k}_t \tilde{x}_t^{\theta-1}. \tag{2A2.24}$$

Steady States

If $(x, \tilde{c}, \tilde{k}_t)$ are steady-state values of $(x_t, \tilde{c}_t, \tilde{k}_t)$ when $(x_t, \gamma_t, n_t, \psi_t)$ are constant at (γ, n, ψ, z), they satisfy:

$$\tilde{c}g[h(z); z] = (1-\theta)h(z)x^{\theta}, \tag{2A2.25}$$

$$1 = \frac{1}{\gamma}\beta \cdot [1 + (1-\tau)\,(\theta x^{\theta-1} - \delta)], \tag{2A2.26}$$

$$\tilde{k} = \frac{1}{\gamma n}\,[(1-\delta) + (1-\psi)x^{\theta-1}]\,\tilde{k} - \tilde{c}. \tag{2A2.27}$$

These three steady-state equations can be solved for $(x, \tilde{c}, \tilde{k})$ as:

$$x = \left(\frac{\dfrac{\frac{\alpha}{\beta}-1}{1-\tau}+\delta}{\theta}\right)^{\frac{1}{\theta-1}}, \tag{2A2.28}$$

$$\tilde{c}_t = (1-\theta)x^{\theta}\frac{h}{g[h(z); z]}, \tag{2A2.29}$$

$$\tilde{k} = \frac{\tilde{c}}{1-\delta+(1-\psi)x^{\theta-1}-\gamma n} = \frac{(1-\theta)x^{\theta}}{1-\delta+(1-\psi)x^{\theta-1}-\gamma n}\cdot\frac{h(z)}{g[h(z); z]}. \tag{2A2.30}$$

So the steady-state value of x, which is pinned down by the Euler equation (2A2.26) alone, is invariant to the functional form of $g(\cdot;\cdot)$. The steady-state value of detrended output $\tilde{y}$ can be calculated easily as:

$$\tilde{y} = \tilde{k}x^{\theta-1} = \frac{(1-\theta)x^{2\theta-1}}{1-\delta+(1-\psi)x^{\theta-1}-\gamma n}\cdot\frac{h(z)}{g[h(z); z]}. \tag{2A2.31}$$

In this expression, as in the two expressions preceding it, x is given by (2A2.28) which does not involve z.

The Level-down Effect

If the steady-state values (γ, n, ψ, z) of the exogenous variables change, then $(x, \tilde{c}, \tilde{k}, \tilde{y})$ change. We say that the *level-down effect* exists if a reduction in

z reduces steady-state output (that is, if $(\partial\tilde{y}/\partial z) > 0$). From the expression for $\tilde{y}$ given in (2A2.31) and the fact that x (the steady-state value of the detrended capital–labor ratio) is independent of z, it is clear that:

$$\frac{\partial\tilde{y}}{\partial z} > 0 \Leftrightarrow \frac{\mathrm{d}\left\{\dfrac{h(z)}{g[h(z);\,z]}\right\}}{\mathrm{d}z} > 0. \qquad (2A2.32)$$

We can check empirically whether this condition for the level-down effect holds, because the first-order condition for e_t, (2A2.8), implies that the ratio $h(z_t)/g[h(z_t);\,z_t]$ should equal the consumption–wage rate ratio $(C_t/N_t)/w_t$. Figure 2A2.2 plots the log of $(C_t/N_t)/w_t$ against z_t (days worked per week) for 1960–2000.[19] It shows, clearly, that the level-down effect existed all through the post-1960 period, except for 1976–84 when the relationship shifted up with the number of days worked at around 5.5.

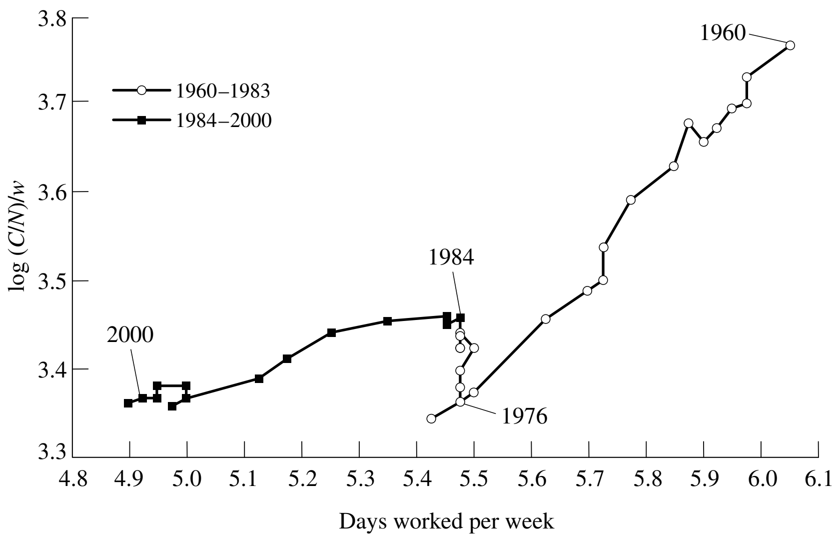


Figure 2A2.2 Log consumption–wage ratio against days worked

Solution Procedure

Returning to the set of equations (2A2.20)–(2A2.22), we now describe our solution procedure for determining the sequence of endogenous variables. Substituting (2A2.20) into (2A2.21) and (2A2.22), the three-equation system can be reduced to a system of two nonlinear difference equations in $(\tilde{c}_t, \tilde{k}_t)$:

$$\tilde{c}_{t+1} = \frac{\tilde{c}_t}{\gamma_t}\beta \cdot \left\{1 + (1-\tau)\left[\theta \cdot \left(\left\{\frac{\tilde{c}_{t+1}}{(1-\theta)\frac{h(z_{t+1})}{g[h(z_{t+1}); z_{t+1}]}}\right\}^{1/\theta}\right)^{\theta-1} - \delta\right]\right\}, \quad (2A2.33)$$

$$\tilde{k}_{t+1} = \frac{1}{\gamma_t n_t}\left[(1-\delta) + (1-\psi_t)\left(\left\{\frac{\tilde{c}_t}{(1-\theta)\frac{h(z_t)}{g[h(z_t); z_t]}}\right\}^{1/\theta}\right)^{\theta-1}\right]\tilde{k}_t - \tilde{c}_t. \quad (2A2.34)$$

This two-equation system generates the sequence of endogenous variables $\{\tilde{c}_t, \tilde{k}_t\}$ given the sequence of exogenous variables $\{\gamma_t, n_t, \psi_t, [h(z_t)]/g[h(z_t); z_t]\}$ as well as calibrated values of the parameters $(\theta, \delta, \beta, \tau)$ and the initial condition for $\tilde{k}_t$. The initial value for $\tilde{k}_t$ (for $t = 1990$) is the actual detrended capital stock at the beginning of 1990. To find the initial value for $\tilde{c}_t$ (for $t = 1990$), a shooting algorithm is used. That is, the value is chosen so that the two-equation system converges to the steady state $(\tilde{c}, \tilde{k})$. With $\{\tilde{c}_t, \tilde{k}_t\}$ thus generated, the sequence of $\{x_t, e_t, \tilde{y}_t\}$ can be calculated from (2A2.20), (2A2.23), and (2A2.24).

The sequence of exogenous variables $\{\gamma_t, n_t, \psi_t\}$ ($t = 1990, 1991, \ldots$) are set as described in the text: γ_t (one plus the growth rate of the TFP factor) is set to its actual value for $t = 1990$–99, and to its average over 1990–99 for $t = 2000$ on; ψ_t (the GNP share of government expenditure) is set to its actual value for $t = 1990$–99, and to its average over 1990–99 for $t = 2000$ on; n_t (one plus the growth rate of working-age population) is its actual value for $t = 1990$–99, and zero for $t = 2000$ on.

Regarding the specification of the sequence of the ratio $[h(z_t)]/g[h(z_t); z_t]$, the most natural idea is to parameterize/calibrate the disutility-of-work function $g(h, z)$, derive the implied function $h(z)$ by (2A2.13), and then generate the sequence of the ratio given the actual sequence of the exogenous variable z_t (days worked). We would base the choice of the function $g(h, z)$ on microeconomic studies of how days worked affect the labor supply. However, such micro studies do not seem to exist in the Japanese literature.

Instead of parameterizing the g function explicitly, we proceed as follows. For $t = 1990, 1991, 1992$, when hours worked were still declining rather sharply (see Figure 2.2 in the text), we assume that the function $g(.; z)$ as a function of h shifted so that the value of x_t (detrended capital–labor ratio) given in (2A2.20) equals the actual detrended capital–labor ratio.

For $t > 1992$, when h_t fluctuated near 40 hours and z_t (days worked) was about 5, we take a linear approximation of $g(h, z)$ around $h = 40$. Since $g(h, z)$ satisfies (2A2.13), the linear approximation has to pass through the origin (see Figure 2A2.1). So the approximated g function must be proportional to h, as in the specification given in equation (2.5) in the text:

$$g(h, z) \approx \frac{\alpha}{40}h. \quad (2A2.35)$$

For this function, the ratio $h/g(h, z)$ equals $40/\alpha$, which is independent of z.

Therefore, for $t = 1990, 1991$, the system of two nonlinear difference equations we use is (2A2.21) and (2A2.22), with the actual detrended capital–labor ratio used for x_{t+1} in the first equation and x_t in the second equation:

$$\tilde{c}_{t+1} = \frac{\tilde{c}_t}{\gamma_t}\beta \cdot [1 + (1-\tau)\,(\theta \bar{x}_{t+1}^{\theta-1} - \delta)],\ \bar{x}_t = \text{\textit{actual} detrended capital–labor ratio,} \tag{2A2.36}$$

$$\tilde{k}_{t+1} = \frac{1}{\gamma_t n_t}[(1-\delta) + (1-\psi_t)\bar{x}_t^{\theta-1}]\tilde{k}_t - \tilde{c}_t,\ \bar{x}_t = \text{\textit{actual} detrended capital–labor ratio.} \tag{2A2.37}$$

For $t = 1993$, the x_{t+1} in the first equation is very close to 40 hours. So the linear approximation of $g(h, z)$ is applicable to the first equation, but not to the second. Thus for $t = 1992$,

$$\tilde{c}_{t+1} = \frac{\tilde{c}_t}{\gamma_t}\beta \cdot \left[1 + (1-\tau)\left(\theta \cdot \left\{\left[\frac{\tilde{c}_{t+1}}{(1-\theta)(40/\alpha)}\right]^{1/\theta}\right\}^{\theta-1} - \delta\right)\right], \tag{2A2.38}$$

$$\tilde{k}_{t+1} = \frac{1}{\gamma_t n_t}[(1-\delta) + (1-\psi_t)\bar{x}_t^{\theta-1}]\tilde{k}_t - \tilde{c}_t,\ \bar{x}_t = \text{\textit{actual} detrended capital–labor ratio.} \tag{2A2.39}$$

For $t > 1993$, the system becomes:

$$\tilde{c}_{t+1} = \frac{\tilde{c}_t}{\gamma_t}\beta \cdot \left[1 + (1-\tau)\left(\theta \cdot \left\{\left[\frac{\tilde{c}_{t+1}}{(1-\theta)(40/\alpha)}\right]^{1/\theta}\right\}^{\theta-1} - \delta\right)\right], \tag{2A2.40}$$

$$\tilde{k}_{t+1} = \frac{1}{\gamma_t n_t}\left((1-\delta) + (1-\psi_t)\left\{\left[\frac{\tilde{c}_t}{(1-\theta)(40/\alpha)}\right]^{1/\theta}\right\}^{\theta-1}\right)\tilde{k}_t - \tilde{c}_t. \tag{2A2.41}$$

Thus, our solution method for determining the sequence of endogenous variables does not require data on z_t and h_t.

To implement this solution procedure, the parameters $(\theta, \delta, \beta, \tau, \alpha)$ need to be calibrated. The calibrated values are the same as in the text: $\theta = 0.362$, $\delta = 0.089$, $\beta = 0.976$, $\tau = 0.480$, $\alpha = 1.373$. Figures 2A2.3–2A2.5 report the simulation obtained from this solution procedure along with the data and the simulation of the text.

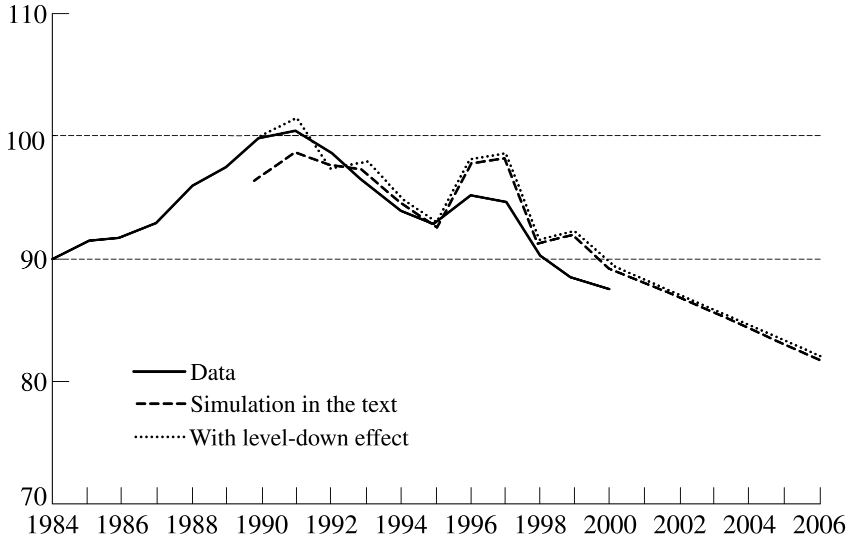


Figure 2A2.3 Detrended real GNP per working-age person (1990 = 100)

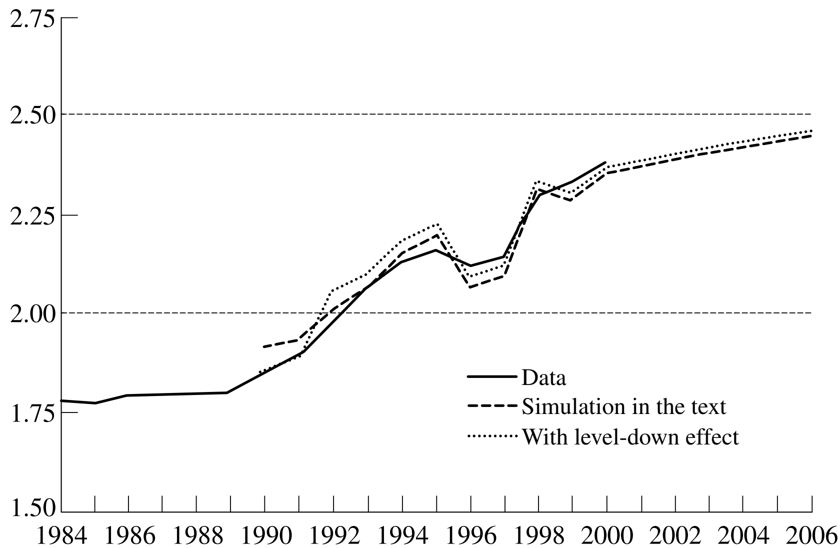


Figure 2A2.4 Capital–output ratio

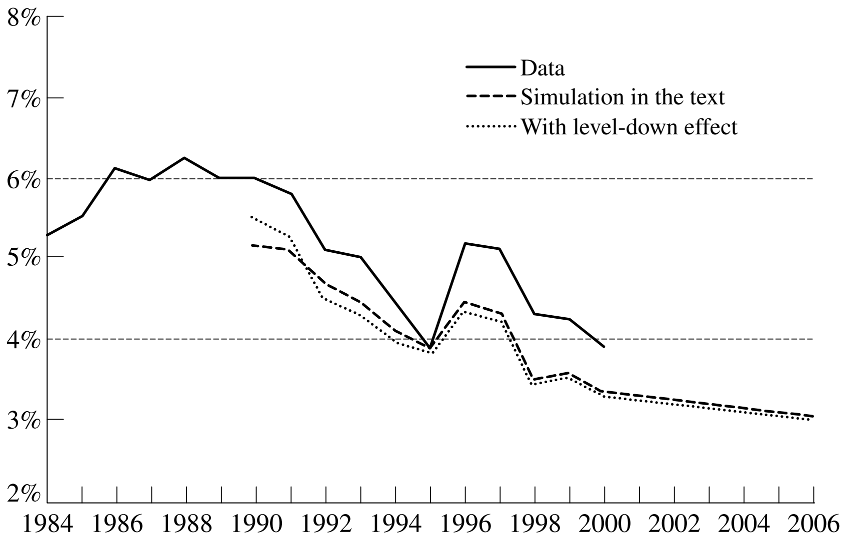


Figure 2A2.5 After-tax rate of return

Relation to the Simulation of the Text

The solution procedure underlying the model prediction in the text uses (2A2.40) and (2A2.41) for all years $t = 1990, 1991, \ldots$, which amounts to assuming no level-down effect.[20] This explains why the decline in GNP is less steep in the text's simulation than in the simulation just described above.

NOTES

* We thank Tim Kehoe, Nobu Kiyotaki, Ellen McGrattan and Lee Ohanian for helpful comments, and Sami Alpanda, Pedro Amaral, Igor Livshits and Tatsuyoshi Okimoto for excellent research assistance and the Cabinet Office of the Japanese government, the United States National Science Foundation, and the College of Liberal Arts of the University of Minnesota for financial support. The views expressed herein are those of the authors.

1. Kwon (1998) and Bayoumi (2001), using VAR analysis, concluded that fluctuations in asset prices affected output through bank lending. Ogawa and Suzuki (1998) find evidence from panel data on large Japanese firms that the price of land as collateral affected investment demand. Sasaki (2000) reports from micro data on Japanese banks that lending by city banks (large Japanese banks) was constrained by the BIS capital ratio requirement. Woo (1999) finds support for the BIS-induced capital crunch only for 1997. Ogawa and Kitasaka (1998, ch. 4) assert that the decline in asset prices shifted both the demand curve and supply curve of bank loans, which resulted in a fall in investment without noticeable change in lending rates. Motonishi and Yoshikawa (1999), while gen-

erally disagreeing with the view that investment was constrained by bank lending, find evidence for a credit crunch for 1997 and 1998.

2. Our procedure for constructing data underlying this and other figures and tables is explained in the Appendix 2A1.
3. The TFP is calculated as $Y_t/(K_t^{\theta}L_t^{1-\theta})$ where the capital share θ is set to 0.362, Y_t is GNP, K_t is the nongovernment capital stock, L_t is aggregate hours worked. The average annual TFP growth rate over 1983–91 is $(A_{1991}/A_{1993})^{1/(1991-1983)} - 1$, which is approximately equal to the average of the annual growth rates between 1984 and 1991.
4. The employer must pay a higher wage rate to have the employee work longer than this statutory limit.
5. The average annual TFP growth rate over 1983–91, for example, is calculated as $(A_{1991}/A_{1993})^{1/(1991-1983)} - 1$.
6. This function is proportional to *h*. See Appendix 2A2 on computation for the reason.
7. Recall that in our accounting framework government investment is included in *G* and that investment consists of domestic private investment and the current account surplus. Hence (2.8) holds with Y_t representing GNP.
8. See Appendix 2A2 for more details of our solution method.
9. See, for example, Kashyap and Stein (1994) for a fuller statement of the hypothesis.
10. Since investment excludes inventory investment here, retention is defined as sales (rather than output) minus the sum of costs, net interest payments, corporate taxes and dividends.
11. The nonfinancial corporate sector in the National Accounts includes public nonfinancial corporations (such as corporations managing subways and airports), which get funding from the Postal Saving System (a huge government bank) through a multitude of government accounts collectively called the Fiscal Investment and Loan Program (FILP). It is not possible to carry out the flow-of-funds analysis for private nonfinancial corporations by excluding public corporations, because the National Accounts do not include a separate capital transactions account for this sector. However, the income-expenditure account for this sector, which is available from the National Accounts, indicates that public nonfinancial corporations are a minor part, less than 10 percent of the nonfinancial corporate sector in terms of income (defined as operating surplus plus property income). Since the nonfinancial corporate sector is the object of our analysis here, the privatization of two large public corporations, Japan Railway and NTT (Nippon Telegraph and Telephone), does not have to be taken into account in our analysis.
12. Bank loans here include loans made by public financial institutions. If loans from public financial institutions are not included, the decline in bank lending in the 1990s is more pronounced.
13. See Appendix 2A1 for more details on this MOF survey.
14. Some of the cash and deposits must be compensating balances. We do not have statistics on compensating balances, however.
15. See Appendix 2A1 for more details.
16. If the loan growth from March 1997 to March 1999 is used instead, the *t* value on loan growth is much smaller.
17. Published data on employment by prefecture are available for Japan, but only for manufacturing and at the ends of calendar years. When we replaced GDP growth by employment growth in the regression, the loan growth coefficient was less significant. For example, if employment growth from December 1996 to December 1998 replaces the GDP growth from 1996–98, the *t* value on the loan growth coefficient is 0.35. Furthermore, in this employment growth equation, if the loan growth is for manufacturing firms, the loan growth coefficient is negative and insignificant.
18. Our view that the credit crunch hypothesis is applicable only for the brief period of late 1997 through early 1998 is in accord with the general conclusion of the literature cited in note 1, particulary Woo (1999) and Motonishi and Yoshikawa (1999).
19. The data source for hours worked per week is the same as that for hours worked. That is, both hours worked and days worked are from the Monthly Labor Survey (*Maitsuki*

Kinro Tokei Chousa) for establishments with 30 or more employees. Those employees include part-time workers. Monthly figures are converted to weekly figures by dividing them by four.

20. This is related to the point made by Aoki et al. (2002) that the model depends on h and e only through the product $h \cdot e$ if the parameterization of the g function is as given in equation (2.5) of the text.

REFERENCES

Aoki, Shuhei, Tomoyuki Ichiba and Minoru Ktiahara (2002), 'A comment: work-week length in Hayashi–Prescott (2002) model', mimeo, November.

Bayoumi, T. (2001), 'The morning after: explaining the slowdown in Japanese growth in the 1990s', *Journal of International Economics*, **53**, April, 241–59.

Bernanke, B. and C. Lown (1991), 'The credit crunch', *Brookings Papers on Economic Activity*, 205–39.

Hansen, G.D. (1985), 'Indivisible labor and the business cycle', *Journal of Monetary Economics*, **16**, 309–27.

Hayashi, F. (1997), *Understanding Savings: Evidence from the United States and Japan*, Cambridge, MA: MIT Press.

Hoshi, T. and A. Kashyap (1999), 'The Japanese banking crisis: where did it come from and how will it end?', *NBER Macro Annual*, National Bureau of Economic Research, 129–201.

Kashyap, A. and J. Stein (1994), 'Monetary policy and bank lending', in G. Mankiw (ed.), *Monetary Policy*, Studies in Business Cycles, vol. 29, Chicago and London: University of Chicago Press, pp. 221–56.

Kwon, E. (1998), 'Monetary policy, land prices, and collateral effects on economic fluctuations. evidence from Japan', *Journal of Japanese and International Economies*, **12**, 175–203.

Motonishi, T. and H. Yoshikawa (1999), 'Causes of the long stagnation of Japan during the 1990s', *Journal of Japanese and International Economies*, **13**, 181–200.

Ogawa, K. and S. Kitasaka (1998), *Asset Markets and Business Cycles* (in Japanese), Tokyo: Nihon Keizai Shinbunsha.

Ogawa, K. and K. Suzuki (1998), 'Land value and corporate investment: evidence from Japanese panel data', *Journal of Japanese and International Economies*, **12**, 232–49.

Peck, M.J., R.C. Levin and A. Goto (1988), 'Picking losers: public policy toward declining industries in Japan', in J.B. Shoven (ed.), *Government Policy Towards Industry in the United States and Japan*, Cambridge: Cambridge University Press, pp. 165–239.

Rogerson, R. (1988), 'Indivisible labor lotteries and equilibrium', *Journal of Monetary Economics,* **21**, 3–16.

Sasaki, Y. (2000), 'Prudential policy for private financial institutions', mimeo, Japanese Postal Savings Research Institute.

Woo, D. (1999), 'In search of "capital crunch": supply factors behind the credit slowdown in Japan', IMF Working Paper No. 99/3, International Monetary Fund, Washington, DC.

3. Effects of information technology and ageing workforce on labor demand and technological progress in Japanese industries: 1980–1998*

Futoshi Kurokawa, Kazunori Minetaki, Kiyohiko G. Nishimura and Masato Shirai

1. INTRODUCTION

The Japanese economy has been in the midst of a fundamental change driven by two powerful forces: the ageing population and the rapid progress in information and communication technology (ICT). The macroeconomic impact of the ageing population on the future of the Japanese economy has attracted much attention, particularly with respect to sustainable pension and health-care systems and the optimal policy mix of debts and taxes to finance government expenditure. Advances in ICT have also been a hot issue, both politically and economically, but most studies have concentrated on the microeconomic impacts of ICT on Japanese firms. These two forces may interact in very important ways: for example, rapid advances in ICT may lead to drastic changes in the structure of jobs and the workplace, which may alter the impacts of the ageing workforce. Few studies have been conducted on these interrelationships and their joint impacts on economic growth.[1] The purpose of this chapter is to fill this gap.

Casual observation suggests that workers and firms can benefit from ICT. The automobile industry has succeeded in substantially shortening the development period of automobiles by relying on computer-aided design (CAD) software. In nonmanufacturing industries, video-scanned data or so-called point-of-sale (POS) systems in the retail sector and automatic teller machines (ATMs) in the banking and securities sectors have decreased costs.

The literature on endogenous growth suggests that technological progress in the workplace may depend on ICT externality. If this dependency is actual and not just theoretical, investment in ICT may play a vital role in sustaining economic growth in Japan, despite the problem of declining numbers of young workers. Some observers claim that the dismal performance of the Japanese economy in the 1990s and early 2000s may stem from insufficient investment in ICT. This view is often supported by comparisons between the revival of the US economy, with its strong ICT investment, and the stagnation of the Japanese economy, with its delayed development of the Internet.[2]

Changes brought about by information technology, however, will not necessarily benefit all workers and firms. Computers help make the compilation and transfer of knowledge easier, and certain skills once acquired from long on-the-job experience may now be shared rapidly among workers, or replaced entirely, with the use of software. Many Japanese workers and firms have relied heavily on nontransferable, relation-specific tacit knowledge concerning how to produce good products and maintain good customer relations. If such knowledge and skills are replaced by software and hardware powered by ICT development, then these workers and firms may lose their comparative advantage in management and production.[3]

Sweeping diffusion of ICT may also have differential impacts in the workplace. Innovations in ICT are likely to demand well-educated, skilled workers as a complementary factor to information- and communication-related capital stocks, and these capital stocks may tend to replace less-educated, unskilled workers. Moreover, if older workers have difficulty in embracing new technology, the overall ageing of the workforce may have a negative impact on productivity growth.

The purpose of this chapter is consequently twofold. First, we examine the direction and the magnitude of substitutability or complementarity between information- and communication-related capital stocks and various labor inputs to discern the differential impacts of ICT on labor demand and to obtain information about the groups of workers for which ICT can effectively substitute. Second, we estimate the contribution of information- and communication-related capital stocks and various labor inputs to the value-added growth of the Japanese economy in the 1980s and 1990s and explore the factors that determine technological progress. In particular, we investigate ICT externality, that is, whether the rapid accumulation of ICT capital stocks has a positive effect on technological progress. We also identify the effects of compositional changes in labor inputs on technological progress, focusing on the ageing workforce and ICT-induced technological obsolescence.

In examining complementarity and substitutability among factors of production, the natural analytical framework is a translog cost function. It is important to note, however, an important caveat in applying such a function to Japanese industries. A translog cost function approach assumes cost minimization coupled with perfect variability of inputs and factor-price-taking behavior. Although the assumption of factor-price-taking behavior reflects the operation of Japanese industries, the perfect-input-variability assumption is problematic. Some factors of production, particularly some categories of capital stocks, are 'quasi-fixed,' that is, they are not completely variable and are fixed in the short run. Buildings and factories are typical examples. Because firms typically have long-term stable relationships with their workers, some part of the workforce can be considered 'quasi-fixed,' and some labor inputs are not sensitive to changes in economic conditions. Personnel in a firm's corporate headquarters can often be considered such quasi-fixed labor inputs, and even production workers often have long-run stable employment relationships with the firms that employ them. In other words, hiring and firing may not be as easy as the translog cost function approach assumes. Thus, it is unclear whether a translog cost function approach is appropriate for analysing the substitutability of all capital stock and labor inputs, although it may be adequate for a subset of these factors.

Our chapter is organized as follows. In Section 2 we present a theory of production capacity functions coupled with capacity utilization functions. The theory explicitly incorporates the quasi-fixed nature of some of the capital stocks and labor inputs, but *it does not presuppose that capital stocks are quasi-fixed and labor inputs are variable*.[4] Whether particular factors are quasi-fixed is an empirical question. We then assume that production capacity is determined by quasi-fixed factors and that capacity utilization is determined by variable factors. Under the assumption of homogeneity of production capacity functions and capacity utilization functions, coupled with long-run constant returns to scale, we show that variable cost shares are independent of output and production capacity, and that basic properties of the variable cost functions can be inferred from the estimated variable cost share functions, without knowledge of output and production capacity. Moreover, we show that this theory allows the estimation of the rate of technological progress without imposing a perfect-competition assumption on product markets. This aspect of the theory is particularly important, because most Japanese industries are not competitive (Ariga et al. 1999; Nishimura et al. 1999).[5] The traditional approach that assumes perfect competition may result in incorrect estimates of technological progress (Nishimura and Shirai 2000).

In Section 3 we construct an industry-wise data set of disaggregated capital stocks and labor inputs. In Section 4 we estimate translog variable cost functions for each industry. We find that the following capital stocks are quasi-fixed: structure, buildings, transportation machines, and machines and tools. In contrast, ICT capital stocks are shown to be variable in all industries. With respect to labor inputs, young workers with low education levels are robustly shown to be variable inputs in all industries. In contrast, old workers with low education levels are robustly shown to be quasi-fixed. The factors that are quasi-fixed and those that are variable are different for the 1980s and the 1990s. The ratio of quasi-fixed factor costs to total costs has increased in the 1990s compared with the 1980s, except in the general machinery industry. This means that Japanese industries were less flexible during the economic turbulence and prolonged downturn of the 1990s. This development may have contributed to the poor performance of Japanese firms during this time.

Three remarkable facts emerge with respect to substitutability and complementarity in capital stocks (to be precise, their services) and labor inputs. First, ICT capital stocks are shown to be significant substitutes for young workers with low education levels, whereas older workers with low education levels are consistently quasi-fixed in all industries under investigation. Second, ICT capital stocks have a complementary relationship with workers with high education levels in many industries. Third, workers with high education levels and those with low education levels are substitutes for one another.

In section 5, we estimate the impact of ICT investment and the changing demography on the rate of technological progress in Japan between 1981 and 1998. The results suggest that the prolonged slump of the 1990s was not merely a demand-driven phenomenon; rather, the supply side played a substantial role. The rate of technological progress declined substantially between the 1980s and 1990s. We examine three possible explanations for the productivity slowdown: (i) inflexibility of older workers in adopting ICT, (ii) technological and managerial obsolescence brought about by ICT development, and (iii) insufficient investment in ICT (failure to realize ICT externality). Initially, our analysis revealed a positive correlation between ICT capital stocks and technological progress in manufacturing, suggesting a strong externality effect of ICT capital stocks, which would support the third explanation for Japan's declining rate of productivity. The correlation is not robust, however: first, if nonmanufacturing industries are included, the correlation vanishes; second, if the electrical machinery sector is excluded from the sample of manufacturing industries, the correlation also vanishes. Thus, we fail to discern clear-cut evidence for ICT externality.[6] The notion that the ICT 'revolution' can boost productivity growth is not supported by our data, although investment in ICT-producing industries is surely an important driving force for economic growth.

Our results do not support the hypothesis that the slowdown in productivity in Japan can be attributed to the inflexibility of the country's older workers. There is no correlation between the rate of technological progress and the ratio of older workers with low education levels to total labor inputs. Our results suggest, however, that ICT development in the 1990s had a negative impact on the past strength of the Japanese economy, a strength often characterized by productivity increases resulting from well-educated workers' learning by doing. In the manufacturing industries in which Japan has traditionally been strong, the rate of technological progress in the 1980s has a positive (though weak) correlation with an older labor force with a high education level. That is, in the 1980s industries having a higher ratio of older well-educated workers to the total labor inputs had a higher rate of technological progress. This suggests that the increased average skill level among well-educated workers resulting from their longer experience has a positive effect on productivity. This relationship changes significantly in the 1990s, however, and the effect becomes negative rather than positive. The nature of technological progress apparently changes adversely in that time period.

We discuss the possible policy implications of our results in section 6. We argue that ICT is an important substitute for young, less-educated workers (who may be increasingly scarce in Japan in the future). The so-called New Economy of ICT externality, however, has not yet materialized. Moreover, the nature of current ICT progress is not favorable to the Japanese economy; that is, ICT externality may have adversely affected the previous comparative advantage of Japanese firms. With these ideas in mind, we point out the necessity of a thorough re-examination of technological policy in Japan.

2. QUASI-FIXED FACTORS, VARIABLE COST FUNCTION, AND MEASUREMENT OF TECHNOLOGICAL CHANGE

Production Function: Production Capacity and Capacity Utilization

Let us consider a production function with n variable factors of production and m quasi-fixed factors:

$$Y = F(x_1, \ldots, x_i, \ldots, x_n; z_1, \ldots, z_j, \ldots, z_m; A)$$

where x_i is the ith variable factor and z_j is the jth quasi-fixed factor. The term A denotes the state of production technology.

Assumption 1 (capacity and utilization) F is multiplicatively separable between variable factors $(x_1, \ldots, x_i, \ldots, x_n)$ and quasi-fixed factors $(z_1, \ldots, z_j, \ldots, z_n)$:

$$Y = F(x_1, \ldots, x_i, \ldots, x_n; z_1, \ldots, z_j, \ldots, z_m; A)$$
$$= G(x_1, \ldots, x_i, \ldots, x_n; A)S(z_1, \ldots, z_j, \ldots, z_m; A) \quad (3.1)$$

$S(z_1, \ldots, z_j, \ldots, z_m; A)$ is the production capacity function, and $G(x_1, \ldots, x_i, \ldots, x_n; A)$ is the capacity utilization function. For example, consider an oil refinery firm. The firm's production capacity is, say, S gallons per day. To realize this capacity, the firm has oil tanks and other large refinery equipment that are fixed in the short run. The firm has maintenance workers and management teams to run a factory of this size. They are also fixed in the short run. The firm produces the actual refinery products by consuming crude oil, servicing trucks and other equipment, and the labor of factory workers.

Assumption 2 (homogeneity) G is homogeneous of degree k in $(x_1, \ldots, x_i, \ldots, x_n)$, and S is homogeneous of degree $1-k$ in $(z_1, \ldots, z_j, \ldots, z_m)$. An immediate consequence of this assumption is that F is homogeneous of degree one in all inputs $(x_1, \ldots, x_i, \ldots, x_n; z_1, \ldots, z_j, \ldots, z_m)$. Thus, we implicitly assume that production exhibits constant returns to scale in the long run where quasi-fixed factors are optimally adjusted.

Assumption 3 (timing of quasi-fixed factor determination) Quasi-fixed factors must be determined one period before production.

Variable Cost Function under the Capacity-cum-utilization Framework

Under assumptions 1 and 2, the share of a variable factor of production in the total variable cost (that is, the variable cost share) is independent of the level of output and production capacity, which means that it can be estimated without knowledge of production capacity. The variable cost function corresponding to the production function F is defined as:

$$C_V = (p_1, \ldots, p_i, \ldots, p_n, Y, S; A)$$
$$= \underset{x_1, \ldots, x_n}{\text{Min}} \sum_{i=1}^{n} p_i x_i \text{ subject to } Y = G(x_1, \ldots, x_i, \ldots, x_n; A)\, S \quad (3.2)$$

With some calculations[7] we have a multiplicatively separable variable cost function, such that:

$$C_V(p_1, \ldots, p_i, \ldots, p_n, Y, S; A) = c_v(p_1, \ldots, p_n; A)\left(\frac{Y}{S}\right)^{1/k}, \quad (3.3)$$

where c_v is homogeneous of degree one in prices. Consequently, using Shepherd's Lemma, we have:

$$\frac{p_i x_i}{C_V} = \frac{p_i}{c_v(p_1, \ldots, p_n; A)} \frac{\partial c_v(p_1, \ldots, p_n; A)}{\partial p_i},$$

which implies that the variable cost share is independent of the level of production (Y) and the level of production capacity (S).

The cost minimization (equation (3.2)) implies:

$$p_i = \lambda \frac{\partial G}{\partial x_i} S \quad \text{for} \quad i = 1, \ldots, n$$

and

$$\lambda = \frac{\partial C_V}{\partial Y}.$$

Moreover, equation (3.3) means that:

$$\frac{Y}{C_V}\frac{\partial C_V}{\partial Y} = \frac{Y}{C_V}\lambda = \frac{1}{k}. \tag{3.4}$$

Using the above result, we have the following formula relating the cost share and the curvature of the production function:

$$\frac{p_i x_i}{C_V} = \frac{1}{C_V}\lambda\frac{\partial G}{\partial x_i}Sxi = \frac{1}{kY}\frac{\partial G}{\partial x_i}Sxi = \frac{1}{kGS}\frac{\partial G}{\partial x_i}Sxi = \frac{1}{k}\left(\frac{x_i}{G}\frac{\partial G}{\partial x_i}\right). \tag{3.5}$$

Quasi-fixed Factor Inputs and Capacity Cost Function

In this section, we explain output and quasi-fixed factor determination for completeness. Output is determined by the gross profit (π_{gross}) maximization:

$$\underset{Y}{\text{Max}}\,\pi_{gross} = p_Y(Y;\Theta)Y - c_v(p_1 \ldots, p_n; A)\left\{\frac{Y}{S}\right\}^{1/k} - J(A),$$

where p_y is the price of output, and Θ denotes other market conditions that determine the competitiveness of the industry in question. The term $J(A) \geq 0$ is the fixed cost that is independent of the quasi-fixed factors. The fixed cost J depends on the state of production technology A. This optimization implies $p_Y = \mu MC$, where:

$$MC \equiv c_v\left(\frac{1}{S}\right)^{1/k}\frac{1}{k}(Y)^{(1/k)-1}$$

and

$$\mu = \mu(\Theta) \equiv \left(1 + \frac{Y}{p_Y}\frac{\partial p_Y}{\partial Y}\right)^{-1}.$$

Here the latter is a markup over marginal cost, MC, which may be different from unity. Thus, we allow imperfect competition in our framework. Then the gross profit is, with some calculation:

$$\pi_{gross} = \pi_{gross}(S) \equiv [(\mu k)^k - 1](\mu k)^{1/(1-k)} p_Y^{\,1/(1-k)} c_v^{\,-k/(1-k)} S^{1/(1-k)}$$

From our assumption of one-period-advance determination of quasi-fixed factors (assumption 3),the quasi-fixed inputs are determined by the following (expected) net profit maximization:

$$\underset{Z_1,\ldots,Z_m}{\text{Max}}\ E_{-1}\textit{net profit}$$

$$\equiv E_{-1}\left\{\pi_{gross}[S(z_1,\ldots,z_j,\ldots,z_m;A)] - \sum_{j=1}^{m} q_{-1,j}z_j - J(A)\right\},$$

where the quasi-fixed factors' prices $\{q_{-1,j}\}$ are those in the previous period, and expectation E_{-1} is calculated using information available in the previous period.[8]

Like the variable-input optimization, the quasi-fixed-input optimization is decomposed into two steps. The first one is the 'capacity cost' minimization. For given S, let the capacity cost C_s be such that:

$$C_s(q_{-1,1},\ldots,q_{-1,j},\ldots,q_{-1,m},S;A) = \text{Min}\sum_{j=1}^{m} q_{-1,j}z_j,$$

subject to:

$$S = S(z_1,\ldots,z_j,\ldots,z_m;A).$$

The optimization implies:

$$q_{-1,j} = \rho\frac{\partial S}{\partial z_j}\ \text{for}\ j = 1,\ldots,m,$$

$$\rho = \frac{\partial C_S}{\partial S},$$

and

$$C_s(q_{-1,1},\ldots,q_{-1,m};A) = c_s\,(q_{-1,1},\ldots,q_{-1,m};A)S^{1/(1-k)}.$$

In the second step, we determine the optimum capacity using this capacity cost function, such that:

$$\underset{S}{\text{Max}}\,E_{-1}\ \textit{net profit} \equiv E_{-1}[\pi_{gross}(S) - c_s(q_{-1,1},\ldots,q_{-1,m};A)S^{1/(1-k)} - J(A)].$$

This maximization determines the optimum capacity S, which in turn determines the quasi-fixed factors.

Note that $S(z_1,\ldots,z_m)$ is homogeneous of degree $1-k$ in $(z_1,\ldots,z_m)$.

Then, using a similar argument to the output elasticity of variable cost, we have the following relationship between capacity cost share and the curvature of the production-capacity function S:

$$\frac{q_{-1,j}z_j}{C_s} = \frac{1}{1-k}\left(\frac{z_j}{S}\frac{\partial S}{\partial z_j}\right). \tag{3.6}$$

Measurement of Technological Progress

We now show that under assumptions 1 and 2, the rate of technological progress can be estimated without making any assumption about the competitiveness of the industries in question.[9]

Let A denote the state of production technology determining production efficiency such that:

$$\begin{aligned} Y &= F(x_1, \ldots, x_j, \ldots, x_n; z_1, \ldots, z_j, \ldots, z_m; A) \\ &= G(x_1, \ldots, x_i, \ldots, x_n; A)S(z_1, \ldots, z_j, \ldots, z_m; A). \end{aligned} \tag{3.7}$$

As usual, we define the rate of technological progress as the output growth that cannot be attributable to factor inputs. Thus, the rate of technological progress (RTP) is defined by:

$$RTP_t = \left[\frac{1}{F}\frac{\partial F}{\partial A}\right]_t \Delta A_t = \left[\frac{1}{GS}\left(\frac{\partial G}{\partial A}S + G\frac{\partial S}{\partial A}\right)\right]_t \Delta A_t.$$

where a suffix t denotes the period, $[X]_t$ is the value of X at the period t, and $\Delta x_t = x_{t+1} - x_t$. With some calculations, we have the following approximate relation:

$$\frac{\Delta Y_t}{Y_t} \approx \sum_{i=1}^{n}\left[\frac{x_i}{G}\frac{\partial G}{\partial x_i}\right]_t \Delta x_{t,i} + \sum_{j=1}^{m}\left[\frac{z_j}{S}\frac{\partial S}{\partial z_j}\right]_t \Delta z_{t,j} + RTP_t.$$

From equations (3.5) and (3.6), we obtain:

$$RTP_t \approx \frac{\Delta Y_t}{Y_t} - k\left(\sum_{i=1}^{n}\frac{p_{t,i}x_{t,i}}{[C_V]_t}\frac{\Delta x_{t,i}}{x_{t,i}}\right) - (1-k)\left(\sum_{i=1}^{m}\frac{q_{t-1,j}z_{t,j}}{[C_S]_t}\frac{\Delta z_{t,j}}{z_{t,j}}\right).$$

Since the variable cost shares and quasi-fixed cost shares are observable, the rate of technological progress is calculated from the above formula if we know k. Thus, the remaining task is to estimate k.

Let us consider the steady state, in which no uncertainty exists. In our framework, the only difference between variable and quasi-fixed inputs is that quasi-fixed inputs must be determined one period before production, when the future is still uncertain. If there is no uncertainty in the future, the sequential optimization described in the previous sections is equivalent to the following one-shot two-step problem. First, for given Y, the steady-state total cost function is defined by:

$$TC^L(p_1, \ldots, p_n, q_{-1,1}, \ldots, q_{-1,m}, Y; A)$$
$$= \underset{x_1, \ldots, x_n, z_1, \ldots, z_m}{\text{Min}} \sum_{i=1}^{n} p_i x_i + \sum_{j=1}^{m} q_{-1,j} z_j \quad \text{s.t.} \quad \text{equation} \qquad (3.7).$$

Then, the optimum steady-state output, Y^L, is determined by:

$$\underset{Y_L}{\text{Max}}\, p_y(Y^L; \Theta)\, Y^L - TC^L(p_1, \ldots, p_n, q_{-1,1}, \ldots, q_{-1,m}, Y^L, A) - J(A).$$

Because the steady-state total cost minimization implies:

$$p_i = \lambda^L \frac{\partial G}{\partial x_i} S \text{ and } q_{-1,j} = \lambda^L G \frac{\partial S}{\partial z_j},$$

we have:

$$C_v^L = \sum_{i=1}^{n} p_i x_i^L = \lambda^L k Y^L \text{ and } C_S^L = \sum_{j=1}^{m} q_{-1,j} z_j^L = \lambda^L (1-k) Y^L.$$

Consequently, we have:

$$k = \frac{C_V^L}{C_V^L + C_S^L} = \frac{C_V^L}{TC^L}.$$

Thus, k is the variable cost's share in the total cost in the steady state of no uncertainty.

If we knew the period in which there was no uncertainty, we could infer k from the variable cost's share of that period. Since we do not a priori know the period of no uncertainty, we approximate k using the time average of the variable cost's share over a relevant period[10] in the empirical analysis of Section 5.

3. DATA: ICT CAPITAL STOCKS AND DISAGGREGATED FACTOR INPUTS

Capital Stocks

ICT capital stocks

We follow Jorgenson and Stiroh (2000) in defining ICT capital stocks. ICT capital stocks consist of ICT hardware and ICT software. ICT hardware includes computer equipment (such as office computers and related instruments) and communication equipment (such as terminal, switching and transmitting devices). The definition of ICT hardware is the same in the United States and Japan. The US definition of ICT software includes prepackaged, custom, and own-account software, but the Japanese definition includes only custom software.[11]

The Ministry of Economy, Trade, and Industry (METI), which was the Ministry of International Trade and Industry (MITI) before January 2001, reports fixed capital formation matrices every five years in the *Base-Year Input Output Tables* (1980, 1985, 1990, 1995), which show the industry-by-industry formation of the above-mentioned disaggregated capital stocks.[12] We further disaggregate these five-year time-aggregate series into annual series by utilizing ICT expenditure data from the *Information Technology Survey* (1980–1998) conducted by the METI.[13] Then ICT capital stock is constructed by applying the perpetual inventory method. For the ICT hardware capital stock series, we use the ICT hardware investment deflators of Schreyer (2000).[14] For the ICT software investment deflator, we assume that the ratio of the hardware deflator to the software deflator is the same in Japan as in the United States, and we construct the software deflator from Schreyer's hardware deflator, relying on the US data reported in Jorgenson and Stiroh (2000).[15]

Structure and (non-ICT) equipment

Miyagawa and Shiraishi (2000) constructed detailed industry-by-industry capital stock series for manufacturing, in which structure capital stocks and equipment capital stocks (including ICT stocks) were separately estimated. Miyagawa et al. (2001) refined this work by including the nonmanufacturing category and by further disaggregating the structure and equipment capital stocks to arrive at five categories of capital stocks: ICT, machines and tools, transportation machines, buildings, and structure.

In the following analysis, we use our estimate of ICT capital stocks (described above) rather than Miyagawa's. There are two reasons for this choice. First, in Miyagawa's classification, ICT capital stocks do not include software stocks. Second, Miyagawa and his associates use the wholesale price index of ICT hardware products (published by the Bank of Japan) as the ICT hardware price deflator. This price index of ICT hardware products (such as computers), however, inadequately separates hardware prices from software prices in both mainframe and personal computers. Thus this wholesale price index does not show a sharp decline of ICT product prices between 1995 and 2000, in stark contrast to the trend observed in the United States. Therefore we adopt Schreyer's index rather than the Bank of Japan's wholesale price index. For other components of capital stocks, we use Miyagawa's series.

As for the estimate of the rental price of these disaggregated capital stocks, we use the following Jorgensonian user-cost formula (except for the investment tax credit, because there is no investment tax credit in Japan):

$$UCC_{it} = \frac{1 - u_t z_{it}}{1 - u_t} (\rho_t + \delta_{it}) q_{it},$$

where UCC_{it} is the user cost of the ith capital stocks, ρ_t is the dividend yield of the Tokyo Stock Exchange, δ_{it} is the ith capital stocks' depreciation rate, u_t is the marginal corporate income tax rate, z_{it} is the ith capital stocks' capital consumption allowance, and q_{it} is the price of the ith capital stocks. The marginal corporate income tax rate, capital consumption allowance and other variables (except for the depreciation rate for ICT stocks) are constructed using the National Tax Agency's *Survey on Corporate Activities*, the Institute of Local Finance's *Annual Statistical Report of Local Governments*, and the Ministry of Finance, Policy Research Institute's *Financial Statements Statistics of Corporations*.[16] As for the depreciation rate of ICT stocks, because we have insufficient data to estimate this rate in Japan, we use the Bureau of Economic Analysis figures for US ICT stocks, which are reported in the *Survey of Current Business* (1997).

Labor Inputs

We construct disaggregated labor input data from a partly unpublished data set of the Ministry of Health, Labor and Welfare's *Basic Survey on Wage Structure*. There are three dimensions in this disaggregation.

Production versus nonproduction workers: manufacturing industries

In manufacturing industries, the *Basic Survey on Wage Structure* distinguishes nonproduction workers from production workers. Production workers are those who engage in operations at production sites, and nonproduction workers are supervisory, clerical and technical workers. (The survey does not include comparable data for the nonmanufacturing industries.)

Young versus old workers: all industries

The *Basic Survey on Wage Structure* includes detailed information about the age of workers in manufacturing and nonmanufacturing industries. We define workers aged 40 years and over as old workers and those aged less than 40 as young workers.

Well-educated versus less-educated workers: all industries

The survey contains only incomplete information on the education levels of workers, so we have only two categories: well-educated workers, defined as those with a college degree or a higher degree (including a degree from a junior college, higher professional school, university or graduate school); and less-educated workers, defined as those with a high school diploma or less (including a diploma from high school, junior high school or elementary school). Disaggregated data on workers' level of education were published for each industry until 1988, when this part of the *Basic Survey on Wage Structure* ceased being published. Fortunately, we were able to obtain post-1989 data from the Ministry of Labor.

Combining the estimated number of employed workers[17] with work-hour data for each industry,[18] we construct labor input data for all industries in the following four categories: (1) young with low education levels, (2) young with high education levels, (3) old with low education levels, and (4) old with high education levels. For the manufacturing sector, we make a finer classification: each of the four categories is divided into (1) production workers and (2) nonproduction workers. Thus, we have eight categories in manufacturing. The hourly wage data for each category are then derived by dividing the compensation data[19] by the total work hours.[20]

Industries under Investigation

We consider both manufacturing and nonmanufacturing industries, using the System of National Accounts (SNA) classification. Table 3.1 shows the industries we consider. We break down manufacturing into 13 industries, following the SNA. Among the 13 industries in the manufacturing sector, we exclude the category 'miscellaneous manufacturing' because it is not a homogeneous industry. The petroleum and coal industry is also excluded, because the data on prices and quantities for this industry are problematic as a result of heavy government intervention and regulations. Thus we consider 11 out of 13 manufacturing industries listed in the SNA classification.

There are seven SNA nonmanufacturing industries, and we exclude the real estate and utilities industries from the estimation. The real estate industry's output includes imputed rents of owner-occupied houses, and these rents are very large in Japan. Thus any changes in real estate output do not properly represent the industry's activities. Utilities are heavily regulated, and because of possible rent-seeking behavior and political influence, this industry may deviate from the cost-minimizing behavior that we postulate in Section 2.[21]

The sample period is 1980–98 for all industries except the finance and insurance industry. The starting year is 1980 because our data sources, especially those for ICT capital stocks, are not always reliable before 1980. The ending year is 1998 because of data availability, except for the finance and insurance industry. For that industry, the ending year is 1995 because

Table 3.1 Industries under study

SNA sector	Abbreviation
Manufacturing industries	
Food and kindred products	Food
Textile mill products	Textile
Paper and allied products	Paper & pulp
Chemicals	Chemicals
Stone, clay, glass	Stone & clay
Primary metal	Pri. metal
Fabricated metal	Fab. metal
Machinery, nonelectrical	Gen. machinery
Electrical machinery	Elec. machinery
Transportation equipment and ordnance	Trans. equipment
Instruments	Instruments
Nonmanufacturing industries	
Construction	Construc.
Trade	Trade
Finance and insurance	Finance
Transportation and communications	Trans. & commu.
Services	Services
Industries excluded from estimation	
Petroleum and coal products	Petro. & coal
Miscellaneous manufacturing	Misc. manufac.
Utilities	Utilities
Real estate	Real estate

problems with nonperforming loans mar the value-added data of this industry after 1995.[22] Descriptive statistics of factor inputs are shown in Table 3.2 (a, b and c) for each industry in two subperiods: 1980–89 and 1990–98.

The tables show that ICT capital stocks increased substantially in almost all industries in the total sample period. Significant differences exist among industries, however, and the industries appear to fall into two groups. The first group, which can be denoted as the 'ICT-intensive industries', shows a rapid accumulation of ICT capital stocks in the 1980s and keeps its lead in the 1990s. There are five ICT-intensive industries: electrical machinery and instruments in the manufacturing sector; and finance and insurance, transportation and communication, and services in the nonmanufacturing sector. These are the same industries that Stiroh (2001) found to be ICT-intensive in

Table 3.2a Average growth rate of factors (percentage): manufacturing

	Food		Textile		Paper & pulp	
	1980s	1990s	1980s	1990s	1980s	1990s
ICT capital	24.43	16.48	22.60	14.08	16.75	10.09
Non-ICT equipment	8.38	3.69	5.90	4.98	4.15	5.10
Structure	2.91	2.45	20.24	20.27	3.52	3.45
Young, low education level	−0.92	−2.11	−4.12	−12.32	−0.49	−1.59
Old, low education level	3.08	20.26	1.47	26.79	2.42	21.47
Young, high education level	2.86	4.28	20.40	22.25	3.59	2.24
Old, high education level	5.44	1.75	5.19	0.98	4.84	3.57
	Chemicals		Stone & clay		Pri. metal	
	1980s	1990s	1980s	1990s	1980s	1990s
ICT capital	28.54	16.88	17.65	15.73	16.96	12.46
Non-ICT equipment	3.63	4.65	6.75	3.15	1.36	3.12
Structure	1.20	2.50	1.43	0.66	1.68	2.48
Young, low education level	−4.05	−2.39	−4.31	−2.24	−2.25	−3.42
Old, low education level	1.51	−0.10	−0.68	−2.32	2.87	−2.75
Young, high education level	1.88	2.34	1.29	1.09	1.63	1.93
Old, high education level	5.44	1.75	5.15	2.66	5.62	1.70
	Fab. metal		Gen. machinery		Elec. machinery	
	1980s	1990s	1980s	1990s	1980s	1990s
ICT capital	30.57	8.12	27.11	11.45	29.11	13.33
Non-ICT equipment	9.68	6.59	10.73	5.20	16.89	7.65
Structure	1.62	2.14	3.02	2.27	8.96	4.03
Young, low education level	−2.78	−2.83	−1.30	−3.33	2.49	−4.93
Old, low education level	0.97	−1.87	2.91	−1.89	6.27	−1.13
Young, high education level	1.78	1.12	3.24	2.21	6.87	0.76
Old, high education level	5.78	3.15	7.02	3.98	8.44	4.21
	Trans. equipment		Instruments			
	1980s	1990s	1980s	1990s		
ICT capital	30.16	12.67	26.56	14.49		
Non-ICT equipment	10.60	5.02	14.20	4.74		
Structure	4.62	2.94	1.73	−0.85		
Young, low education level	−2.16	−2.43	−4.43	−6.18		
Old, low education level	1.50	−1.18	0.56	−1.80		
Young, high education level	2.61	2.47	2.81	−1.57		
Old, high education level	4.90	5.58	5.11	4.12		

Notes: Average growth rates are annual percentage rates. 1980s = 1981–89, 1990s = 1990–98. Young = under 40. Old = 40 and over. Low education level = with high school or less education. High education level = with junior college or more education. See Table 3.1 for industry abbreviations.

Table 3.2b Average growth rate of factors (percentage): nonmanufacturing

	Construc.		Trade		Finance	
	1980s	1990s	1980s	1990s	1980s	1990s
ICT capital	12.04	14.77	14.94	11.81	24.77	n.a.
Non-ICT equipment	3.66	1.65	6.97	2.34	8.44	n.a.
Structure	3.88	4.19	3.96	4.32	2.36	n.a.
Young, low education level	−2.41	−0.71	−3.00	−4.20	24.02	n.a.
Old, low education level	0.83	−20.33	2.34	−0.26	2.00	n.a.
Young, high education level	1.70	2.42	3.68	1.20	5.69	n.a.
Old, high education level	6.48	5.70	7.15	4.37	7.02	n.a.

	Trans. & commu.		Services	
	1980s	1990s	1980s	1990s
ICT capital	12.38	12.67	20.35	18.15
Non-ICT equipment	8.55	2.11	8.95	6.56
Structure	3.30	3.85	7.14	6.06
Young, low education level	−1.92	−1.83	0.44	−0.87
Old, low education level	3.27	−0.16	3.21	0.58
Young, high education level	2.85	5.94	6.04	4.60
Old, high education level	6.23	5.43	7.35	6.53

Notes: Average growth rates are annual percentage rates. 1980s = 1981–89. 1990s = 1990–98. Young = under 40. Old = 40 and over. Low education level = with high school or less education. High education level = with junior college or more education. See Table 3.1 for industry abbreviations.

the United States.[23] The remaining 11 industries are ICT-nonintensive industries. For other types of capital stocks, there is no such salient difference among industries.

For labor inputs, differences among various industries are wide. In general, however, over time the percentage of old workers in the total workforce, the percentage of well-educated workers in the total workforce, and the ratio of production workers to nonproduction workers in the case of manufacturing, show little change within the various industries over the study period.

Table 3.2c Average growth rate of factors (percentage): manufacturing, production and nonproduction labor

	Food		Textile		Paper & pulp	
	1980s	1990s	1980s	1990s	1980s	1990s
Production, low education level	1.33	−0.73	−1.04	−9.18	0.83	−1.36
Production, high education level	6.31	6.95	6.59	−0.38	6.80	6.90
Nonproduction, young, low education level	−1.71	−3.90	−5.10	−9.43	−0.32	−2.36
Nonproduction, old, low education level	4.06	−0.02	0.49	−5.36	3.81	−1.97
Nonproduction, young, high education level	2.24	3.90	−1.80	−2.24	3.23	1.05
Nonproduction, old, high education level	4.96	5.68	3.36	0.14	3.93	2.45
	Chemicals		**Stone & clay**		**Pri. metal**	
	1980s	1990s	1980s	1990s	1980s	1990s
Production, low education level	−1.38	−0.57	−2.22	−2.51	0.71	−2.98
Production, high education level	2.42	7.97	4.01	7.58	6.23	4.41
Nonproduction, young, low education level	−5.45	−4.89	−4.27	−3.16	−2.88	−4.04
Nonproduction, old, low education level	2.63	20.08	0.62	−0.42	2.75	−2.32
Nonproduction, young, high education level	1.98	1.67	1.03	−0.10	1.62	1.28
Nonproduction, old, high education level	5.29	1.14	4.92	1.22	4.57	1.74
	Fab. metal		**Gen. machinery**		**Elec. machinery**	
	1980s	1990s	1980s	1990s	1980s	1990s
Production, low education level	−1.00	−2.50	0.96	−2.82	3.83	−3.43
Production, high education level	4.29	4.96	7.94	5.03	7.91	2.47
Nonproduction, young, low education level	−3.08	−2.12	−2.54	−3.42	1.43	−5.41

Table 3.2c (*continued*)

	Fab. metal		Gen. machinery		Elec. machinery	
	1980s	1990s	1980s	1990s	1980s	1990s
Nonproduction, old, low education level	2.57	−1.22	3.38	−0.85	8.23	−0.52
Nonproduction, young, high education level	1.60	0.37	2.62	1.88	6.82	0.72
Nonproduction, old, high education level	5.16	2.25	6.57	3.31	8.20	3.83
	Trans. equipment		Instruments			
	1980s	1990s	1980s	1990s		
Production, low education level	−0.47	−2.20	−3.34	−4.03		
Production, high education level	6.14	6.85	20.36	3.01		
Nonproduction, young, low education level	−3.54	−2.11	−2.70	−7.24		
Nonproduction, old, low education level	2.42	0.67	2.38	−1.34		
Nonproduction, young, high education level	2.34	2.03	3.59	−2.16		
Nonproduction, old, high education level	4.47	4.90	5.06	3.78		

Notes: Average growth rates are annual percentage rates. 1980s = 1981–89. 1990s = 1990–98. Young = under 40. Old = 40 and over. Low education level = with high school or less education. High education level = with junior college or more education. Production = engaging in operations at production sites. Nonproduction 5 engaging in supervisory, clerical, and technical work. See Table 3.1 for industry abbreviations.

4. SUBSTITUTABILITY BETWEEN ICT CAPITAL STOCKS AND LABOR INPUTS: 1980–98

In this section, we examine the impact of advances in ICT on the demand for labor inputs. In particular, we explore whether ICT capital stocks are substitutes or complements with respect to various labor inputs and whether the magnitude of such substitutability or complementarity has changed between the 1980s and 1990s.[24]

To examine substitutability or complementarity, we have to determine which factors are variable and which factors are quasi-fixed. If all factors are variable, cost minimization implies that their input prices are equal to their marginal product, and this information is utilized to calculate substitutability or complementarity by estimating cost functions. If some factors are variable but others are quasi-fixed, however, the standard procedure cannot be applied, because the input prices of the quasi-fixed factors are no longer equal to their marginal product.

As explained in Section 1, in the context of the Japanese economy it is not appropriate to assume a priori that all factor inputs are variable or that capital stocks are quasi-fixed and labor inputs are variable. Rather, using the framework of Section 2, we let the data determine which factor inputs are variable.

Translog Variable Cost Function under the Capacity-cum-utilization Framework

Since our concern is to detect substitutability or complementarity between factor inputs, we should use a flexible function form for cost functions, allowing both substitutability and complementarity. For this purpose, we use a translog cost function. Let n be the number of variable inputs. We assume that c_v in equation (3.3) has a translog functional form such that:

$$\log c_v(p_1, \ldots, p_n; A) = \alpha(A) + \sum_{i=1}^{n} \beta_i(A) \log p_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij}(A) bgp_i \log p_{jg},$$

In order for C_v to be a cost function, c_v should be nondecreasing and homogeneous of degree one in $(p_1, \ldots, p_n)$. It can be shown that the following restrictions on parameters of c_v are sufficient to satisfy these requirements:[25]

$$\sum_{i=1}^{n} \beta_i(A) = 1, \sum_{i=1}^{n} \gamma_{ij}(A) = 0, \sum_{j=1}^{n} \gamma_{ij}(A) = 0.$$

Under these restrictions, we immediately get the cost share function such that:

$$\frac{p_i x_i}{C_V} = \beta_i(A) + \sum_{j=2}^{n} \gamma_{ij}(A)(\log p_j - \log p_1),$$

for $i = 2, \ldots, n$, which can be estimated using information about variable factor shares and variable input prices.

One remaining requirement on C_V is that C_V should be concave in $(p_1, \ldots, p_n)$.[26] This requirement is satisfied if and only if c_v is concave. It should be noted that the concavity property of cost functions depends on the assumption that the decision maker can freely choose factor inputs and minimizes costs by appropriately adjusting factor inputs to their price changes. That is, the concavity is a necessary condition for factors ($x_1, \ldots, x_n$) to be variable. Thus, if some factors are fixed in the short run, then an estimated cost function assuming these factors to be variable may not exhibit the concavity property. In other words, if an estimated cost function fails to exhibit the concavity property, then some factors may be (quasi-)fixed rather than variable. We use this property to determine whether a particular factor input is quasi-fixed.

Let us now consider the effect of a change in production technology, which is represented by a change in A. As is well known, Hicks-neutral technological progress (A_H), such that $G(x_1, \ldots, x_i, \ldots, x_n; A_H) = A_H G^*(x_1, \ldots, x_i, \ldots, x_n)$ for some G^*, does not affect β_i and γ_{ij}. However, non-Hicks-neutral technological progress may change either or both of β_i and γ_{ij}. There is no a priori reason to assume that technological change is Hicks neutral, so we should take the effect of the change in A on $\beta_i(A)$ and $\gamma_{ij}(A)$ explicitly into consideration.

The sample period, 1980–98, was very turbulent. In the 1980s the Japanese economy was in its heyday, with good economic growth, booming asset markets and no inflation. In the 1990s, Japan experienced stagnation and depressed asset markets. Thus, the composition of industry products may differ and consequently value-added production technology may differ between the two decades. Moreover, information technology has an effect on the workplace. Two waves of ICT innovation occurred in the economy: the first was in the mid-1980s, and the second was in the 1990s. Typical examples of ICT innovation in the 1980s are factory automation and CAD in manufacturing, particularly in the machinery industries, POS systems in the wholesale and retail trade industry, and ATMs and cash dispensers in the finance and insurance industry. In the early 1990s, personal computers became widely used in offices, and Internet use became popular after 1993. These ICT innovations allegedly changed so-called 'white-collar' jobs. Moreover, customer relationship management and supply chain management software was introduced and had a great impact on the organization of many firms. These developments suggest that value-added production technology may have changed between the 1980s and the 1990s.

In order to take account of possible non-Hicks-neutral changes in production technology, we employ a period-dummy framework such that:

$$\frac{p_i x_i}{C_V} = \left(\beta_i + \sum_k \beta^I_{i,s_k} D^I_{i,s_k}\right) + \left[\sum_{j=2}^{n} \left(\gamma_{ij} + \sum_k \gamma^S_{ij,s_k} D^S_{ij,s_k}\right)\right] (\log p_j - \log p_1), \tag{3.8}$$

for $i = 2 \ldots n$. Here D^I_{i,s_k} is the intercept period dummy and D^S_{ij,s_k} is the slope period dummy, representing effects of (non-Hicks-neutral) technological change in period s_k, which are 0 before s_k, and 1 after s_k. The symmetry of γ_{ij} is imposed in estimation.

Identification of Variable Factor Inputs and Estimation of Variable Cost Parameters: A Heuristic Approach

As explained in the previous section, we have five categories of capital stocks (ICT, machines and tools, transportation machines, buildings, and structure) and four categories of labor inputs (young with low education, young with high education, old with low education, and old with high education). We further decompose the manufacturing sector into two subcategories: nonproduction and production. Thus, many possible input combinations must be tested to determine whether the estimated cost functions that use these inputs will satisfy the monotonicity and concavity requirements. The results of these tests will allow us to determine which factors can be considered variable inputs. Our sample presents some problems. The data are annual, and our time period is only 19 years. Given that the differences among industries are expected to be large, a simple pooling of all industries to cope with the small-sample problem may not be appropriate. This problem is compounded by the possibility that technological change over our sample period may have affected the factor inputs, so that some may have changed from variable to quasi-fixed, or vice versa.

We cope with these problems heuristically. As a starting point, we postulate the order of 'quasi-fixedness'. With respect to capital stocks, we assume that structure capital (structure and buildings) is more likely to be quasi-fixed than non-ICT equipment (machines and tools and transportation machines – mostly automobiles) and that non-ICT equipment is more likely to be quasi-fixed than ICT capital stocks. Our characterization of the mobility or quasi-fixedness of these capital stocks is consistent with Fraumeni (1997), who reports that the service life of ICT capital stocks is roughly 3–5 years, that of machines and tools and automobiles is 8–15 years, and that of structure capital is more than 20 years. As for labor inputs, we postulate that young workers with low levels of education are less likely to be quasi-fixed than other workers. We also assume that young workers with low education are the least 'quasi-fixed' category among all factor inputs except for ICT capital stocks. We make no other assumptions about the relative quasi-fixedness of labor inputs. We choose ICT capital stocks as factor 1 of the regression equation (3.8). This turns out to be satisfactory, because the estimated cost functions (of some form) with ICT capital stocks as factor 1 exhibit concavity in all industries. We employ the following three-step procedure industry by industry. More details about this procedure are given in Appendix 3A3.

In step 1, we choose variable inputs. For all industries, we have two labor input dimensions (young/old and high education/low education) and thus four types of labor inputs. This is the baseline case. We estimate equation (3.8) and determine whether the estimated cost function parameters are consistent with the concavity requirement. In this step, we ignore technological change and thus estimate equation (3.8) without period dummies. The estimation method utilizes full-information maximum likelihood, and the estimation period is from 1980 to 1998. Because we ignore technological change, we do not expect sharp results but we should get reasonably good results in which all estimated γ_{ii}'s are negative, at least with marginal statistical significance, to satisfy the concavity requirement, although they may not be significant under standard levels of significance (5 percent or 1 percent). We apply this procedure starting from five factor inputs, and if the result is unsatisfactory, we drop some of the factors until we get satisfactory results (see Appendix 3A3). In the end, we are able to identify variable inputs for all industries. Moreover, the list of variable factors is unique to each industry; that is, only one combination of factor inputs is chosen for each industry by this procedure.

In step 2, we re-estimate equation (3.8) with period dummies that represent technological change. As explained above, ICT may have affected factor inputs differently in the 1980s from in the 1990s. Thus, we consider both the intercept and slope dummies of the 1990s. Additional industry-specific technological changes may have occurred. In fact, in some manufacturing industries, the contribution of ICT capital stocks changed abruptly in the late 1990s, which may indicate that another technological change occurred later in the 1990s. In the case of transportation and communication, a sharp increase in the ICT contribution is found in the 1980s. This suggests that a technological change may have occurred in the mid-1980s in this industry. Taking these observations into account, we consider additional intercept and slope dummies of the mid-1990s for manufacturing industries having a break in the ICT contribution in the 1990s. In the case of transportation and communication, we consider the mid-1980s dummies instead of the mid-1990s dummies.

After determining the number of period dummies (that is, technological changes), we estimate equation (3.8) with these period dummies. We drop insignificant intercept and/or slope dummies, re-estimate the equations, and determine whether the estimated coefficients are consistent with the concavity requirement. With respect to the timing of technological change, we use the following heuristic search procedure. For possible technological change from the 1980s to the 1990s, we initially assume that the change takes place in 1990. For possible technological change in the mid-1990s (or mid-1980s), we assume that the change takes place in 1995 (or 1985). We then move the point of change around the initial point to see whether this gives us a sharper estimation (in terms of statistical significance of γ_{ii}), with the concavity requirement remaining satisfied. If a particular specification yields a sharper result, we choose that specification; otherwise, we stick to the original specification.

Step 3 is applied only to the manufacturing industries. For these industries, we have one more dimension with respect to labor input types: production workers and nonproduction workers. Production workers are factory workers engaging in production, whereas nonproduction workers are supervisory, clerical and technical workers. We do not group these two types of workers together because their activities differ, and this might lead to misleading results. Basically, we repeat steps 1 and 2 using these additional categories for the labor input data of manufacturing. A problem arises, however, with apparent multicollinearity. To avoid this problem, we aggregate young and old workers in the production worker category. Thus, we consider six types of labor inputs (production workers with low levels of education, production workers with high levels of education, nonproduction young workers with low levels of education, nonproduction young workers with high levels of education, nonproduction old workers with low levels of education, and nonproduction young workers with high levels of education).

Age, Education Level and ICT Stocks

Tables 3.3–3.8 report the results of the procedure outlined above. Table 3.3a shows the periods of (non-Hicks-neutral) technological change detected in the baseline case of four labor input types. In most baseline cases, the concavity requirement is satisfied at the average input price of the relevant period.[27] In a few industries, the concavity requirement is not strictly satisfied for some subperiods, but the deviation from the requirement is rather small.[28] Table 3.3b shows the result for the extended case of six labor inputs in manufacturing industries. Because of the apparent multicollinearity problem explained earlier, we get satisfactory results for only four out of 11 manufacturing industries.

Tables 3.4a and 3.4b list which factor inputs are variable in each industry. As expected, ICT capital stocks and young workers with low education levels are shown to be variable in all industries. In the case of manufacturing industries, ICT capital stocks and production workers with low levels of education are variable when the production versus nonproduction difference is explicitly considered (Table 3.4b). The share equation estimation

Table 3.3a Technological changes and concavity of cost functions for baseline case (four types of labor input)

Industry	Period(s) in which technological change is detected		
Manufacturing			
Food	80–89 OK	90–98 OK	
Textile	80–92 OK	93–98 OK	
Paper & pulp	80–88 OK	89–92 OK*	93–98 OK*
Chemicals	80–89 OK	90–98 OK	
Stone & clay	80–91 OK	92–98 OK	
Pri. metal	80–87 OK	88–94 OK	95–98 OK
Fab. metal	80–87 OK	88–93 OK	94–98 OK
Gen. machinery	80–89 OK	90–98 OK	
Elec. machinery	80–92 OK	93–98 OK	
Trans. equipment	80–88 OK*	89–92 OK*	93–98 OK
Instruments	80–93 OK	94–98 OK	
Nonmanufacturing			
Construc.	80–89 OK	90–98 OK	
Trade	80–98 OK		
Finance (only before 1995)	80–95 OK		
Trans. & commu.	80–85 OK	86–89 OK	90–98 OK
Services	80–89 OK	90–98 OK	

Notes: Concavity is evaluated at the average input prices. OK = sufficient conditions of strict concavity are satisfied. OK* = conditions are not strictly satisfied, but the deviations from them are negligible. See Table 3.1 for industry abbreviations.

Table 3.3b Technological changes and concavity of cost functions for extended case (six types of labor input)

Industry	Period(s) in which technological change is detected		
Manufacturing			
Paper & pulp	80–88 OK	89–92 OK	93–98 OK
Stone & clay	80–84 OK	85–91 OK	92–98 OK
Elec. machinery	80–85 OK	86–88 OK	89–98 OK
Trans. equipment	80–87 OK	88–89 OK	90–98 OK*

Notes: OK = sufficient conditions of strict concavity are satisfied. OK* = conditions are not strictly satisfied, but deviations from them are negligible. Six series labor data can be available only in manufacturing. See Table 3.1 for industry abbreviations.

Table 3.4a List of variable factor inputs by industry: all industries (four types of labor input)

Industry	Types of variable input	
Industries with three variable factor inputs		
Food	IT	IT capital stocks
	YL	Young worker with low education level
	YH	Young worker with high education level
Textile	IT	IT capital stocks
	YL	Young worker with low education level
	YH	Young worker with high education level
Fab. metal	IT	IT capital stocks
	YL	Young worker with low education level
	YH	Young worker with high education level
Gen. machinery (1990s)	IT	IT capital stocks
	YL	Young worker with low education level
	YH	Young worker with high education level
Elec. machinery	IT	IT capital stocks
	YL	Young worker with low education level
	YH	Young worker with high education level
Instruments	IT	IT capital stocks
	YL	Young worker with low education level
	OH	Old worker with high education level
Services	IT	IT capital stocks
	YL	Young worker with low education level
	OH	Old worker with high education level
Industries with two variable factor inputs		
Paper & pulp	IT	IT capital stocks
	YL	Young worker with low education level
Chemicals	IT	IT capital stocks
	YL	Young worker with low education level
Stone & clay	IT	IT capital stocks
	YL	Young worker with low education level
Pri. metal	IT	IT capital stocks
	YL	Young worker with low education level
Gen. machinery (1980s)	IT	IT capital stocks
	YL	Young worker with low education level
Trans. equipment	IT	IT capital stocks
	YL	Young worker with low education level
Construction	IT	IT capital stocks
	YL	Young worker with low education level

Table 3.4a (*continued*)

Trade	IT YL	IT capital stocks Young worker with low education level
Finance (before 1995)	IT YL	IT capital stocks Young worker with low education level
Trans. & commu.	IT YL	IT capital stocks Young worker with low education level

Notes: Young = under 40. Old = 40 and over. Low education level = with high school or less education. High education level = with junior college or more education. See Table 3.1 for industry abbreviations.

Table 3.4b List of variable factor inputs by industry: manufacturing (six types of labor input)

Industry	Types of variable inputs	
Industries with three variable factor inputs		
Elec. machinery	IT PL NPYH	IT capital stocks Production worker with low education level Nonproduction young worker with high education level
Trans. equipment	IT PL NPOH	IT capital stocks Production worker with low education level Nonproduction old worker with high education level
Industries with two variable factor inputs		
Paper & pulp	IT PL	IT capital stocks Production worker with low education level
Stone & clay	IT PL	IT capital stocks Production worker with low education level

Notes: Young = under 40. Old = 40 and over. Low education level = with high school or less education. High education level = with junior college or more education. Production = engages in operations at production sites. Nonproduction = engages in supervisory, clerical and technical work. See Table 3.1 for industry abbreviations.

shows that the coefficients of the estimated share equations are statistically significant and the equations have high adjusted R^2 values and no sign of autocorrelation (see Tables 3.5a and 3.5b). One exception is the electrical machinery industry in the baseline case of four labor input types, which shows a strong error autocorrelation. In the extended case of the six labor inputs in the manufacturing industries, this error autocorrelation is reduced

Table 3.5a Estimation of cost share functions: all industries (four types of labor input)

Manufacturing

Food (Period: 1980–98) 1 = ICT, 2 = YL, 3 = YH

	Coefficient	*t*-statistic	Prob.
β_2	0.836835	86.47719	0.0000
β_3	0.159113	26.64262	0.0000
γ_{22}	−0.190435	−18.77942	0.0000
γ_{23}	0.078321	11.62105	0.0000
1990 DUMMY for β_2	−0.045329	−4.22654	0.0002
1990 DUMMY for γ_{23}	0.012446	3.157459	0.0035
Share Equation YL			
Adjusted R-square		0.965777	
Q-statistics (H_0: no auto-correlation)		0.0907	(p-value: 0.763)
Share Equation YH			
Adjusted R-square		0.957404	
Q-statistics (H_0: no auto-correlation)		0.000006	(p-value: 0.994)

Textile (Period: 1980–98) 1 = ICT, 2 = YL, 3 = YH

	Coefficient	*t*-statistic	Prob.
β_2	0.860119	97.78446	0.0000
β_3	0.117813	13.52187	0.0000
γ_{22}	−0.207041	−12.93283	0.0000
γ_{23}	0.079255	9.819605	0.0000
1993 DUMMY for β_3	0.010277	1.816102	0.0784
Share Equation YL			
Adjusted R-square		0.931533	
Q-statistics (H_0: no auto-correlation)		1.96651	(p-value: 0.161)
Share Equation YH			
Adjusted R-square		0.921065	
Q-statistics (H: no auto-correlation)		0.6695	(p-value: 0.413)

Table 3.5a (*continued*)

Paper & pulp (Period: 1980–98)	1 = ICT, 2 = YL		
	Coefficient	*t*-statistic	Prob.
β_2	0.994474	197.9274	0.0000
γ_{22}	−0.02238	−5.825194	0.0000
1989 DUMMY for β_2	−0.014646	−5.678358	0.0000
1993 DUMMY for γ_{22}	0.003746	2.548571	0.0223
Share Equation YL			
Adjusted R-square		0.945082	
Q-statistics (H_0: no auto-correlation)		0.4674	(p-value: 0.494)

Chemical (Period: 1980–98)	1 = ICT, 2 = YL		
	Coefficient	*t*-statistic	Prob.
β_2	1.168784	78.76967	0.0000
γ_{22}	−0.147726	−17.39529	0.0000
1990 DUMMY for β_2	0.133694	2.668878	0.0175
1990 DUMMY for γ_{22}	−0.08189	−3.937875	0.0013
Share Equation YL			
Adjusted R-square		0.989087	
Q-statistics (H_0: no auto-correlation)		0.3095	(p-value: 0.578)

Stone & Clay (Period: 1980–98)	1 = ICT, 2 = YL		
	Coefficient	*t*-statistic	Prob.
β_2	1.035082	164.6683	0.0000
γ_{22}	−0.036697	−7.748249	0.0000
1992 DUMMY for γ_{22}	0.006052	2.937864	0.0097
Share Equation YL			
Adjusted R-square			0.915275
Q-statistics (H_0: no auto-correlation)		0.467	(p-value: 0.494)

Pri. metal (Period: 1980–98)	1 = ICT, 2 = YL		
	Coefficient	*t*-statistic	Prob.
β_2	1.090819	65.21458	0.0000
γ_{22}	−0.097144	−8.389374	0.0000
1988 DUMMY for β_2	−0.02254	−2.898441	0.0110
1995 DUMMY for γ_{22}	0.017	5.358759	0.0001
Share Equation YL			
Adjusted R-square		0.95987	
Q-statistics (H_0: no auto-correlation)		1.1795	(p-value: 0.277)

Fab. metal (Period: 1980–98)	1 = ICT, 2 = YL, 3 = YH		
	Coefficient	*t*-statistic	Prob.
β_2	0.866233	80.60727	0.0000
β_3	0.131824	7.587936	0.0000
γ_{22}	−0.148852	−7.391223	0.0000
γ_{23}	0.064229	5.091403	0.0000
1988 DUMMY for β_2	−0.01751	−2.884108	0.0070
1994 DUMMY for β_3	0.016034	4.75823	0.0000
Share Equation YL			
Adjusted R-square		0.951908	
Q-statistics (H_0: no auto-correlation)		1.1816	(p-value: 0.277)
Share Equation YH			
Adjusted R-square		0.923458	
Q-statistics (H_0: no auto-correlation)		0.7403	(p-value: 0.39)

Gen. machinery (Period: 1980–89)	1 = ICT, 2 = YL		
	Coefficient	*t*-statistic	Prob.
β_2	1.043333	−63.9369	0.0000
γ_{22}	−0.051775	−15.36908	0.0000
Share Equation YL			
Adjusted R-square		0.905068	
Q-statistics (H_0: no auto-correlation)		1.1252	(p-value: 0.289)

Table 3.5a (*continued*)

Gen. machinery (Period: 1990–98)	1 − ICT, 2 − YL, 3 − YH		
	Coefficient	t-statistic	Prob.
β_2	0.751362	45.29617	0.0000
β_3	0.190252	12.64918	0.0000
γ_{22}	−0.400227	−49.81038	0.0000
γ_{23}	0.301408	176.1062	0.0000
γ_{33}	−0.197548	−26.04773	0.0000
Share Equation YL			
Adjusted R-square		0.884776	
Q-statistics (H_0: no auto-correlation)		0.8353	(p-value: 0.361)
Share Equation YH)			
Adjusted R-square		0.86635	
Q-statistics (H_0: no auto-correlation)		1.2182	(p-value: 0.270)

Elec. machinery (Period: 1980–98)	1 = ICT, 2 = YL, 3 = YH		
	Coefficient	t-statistic	Prob.
β_2	0.870574	39.05199	0.0000
β_3	0.180672	9.015999	0.0000
γ_{22}	−0.43485	−8.302452	0.0000
γ_{23}	0.207779	3.885117	0.0005
γ_{33}	−0.111772	−2.036046	0.0501
1993 DUMMY for γ_{22}	0.027052	3.911322	0.0004
Share Equation YL			
Adjusted R-square		0.96984	
Q-statistics (H_0: no auto-correlation)		5.8219	(p-value: 0.016)
Share Equation YH			
Adjusted R-square		0.95655	
Q-statistics (H_0: no auto-correlation)		3.3269	(p-value: 0.068)

Trans. equipment (Period: 1990–98)		1 = ICT, 2 = YL	
	Coefficient	*t*-statistic	Prob.
β_2	1.021142	103.2854	0.0000
γ_{22}	−0.037567	−5.963609	0.0000
1989 DUMMY for β_2	−0.020381	−3.91411	0.0014
1993 DUMMY for γ_{22}	0.004731	2.354335	0.0326
Share Equation YL			
Adjusted R-square		0.942919	
Q-statistics (H_0: no auto-correlation)		1.7248	(p-value: 0.189)

Instruments (Period: 1980–98)		1 = ICT, 2 = YL, 3 = OH	
	Coefficient	*t*-statistic	Prob.
β_2	0.61044	3.862644	0.0005
β_3	0.299812	1.873122	0.0702
γ_{22}	−0.655653	−3.735925	0.0007
γ_{23}	0.438882	2.50539	0.0175
γ_{33}	−0.301693	−1.707445	0.0974
1993 DUMMY for γ_{22}	0.011972	3.369488	0.0020
Share Equation YL			
Adjusted R-square		0.959192	
Q-statistics (H_0: no auto-correlation)		0.9363	(p-value: 0.333)
Share Equation OH			
Adjusted R-square		0.96144	
Q-statistics (H_0: no auto-correlation)		0.3987	(p-value: 0.528)

Nonmanufacturing

Construc. (Period: 1980–98)		1 = ICT, 2 = YL	
	Coefficient	*t*-statistic	Prob.
	$\beta_2$1.029154	196.0772	0.0000
γ_{22}	−0.019982	−6.518501	0.0000
1990 DUMMY for β_2	−0.003151	−2.803711	0.0134
1993 DUMMY for γ_{22}	0.001269	3.283372	0.0050

Table 3.5a (*continued*)

Share Equation YL			
Adjusted R-square		0.975001	
Q-statistics (H_0: no auto-correlation)		0.6697	(p-value: 0.413)

Trade (Period: 1980–98)		1 = ICT, 2 = YL	
	Coefficient	t-statistic	Prob.
β_2	1.065306	172.4378	0.0000
γ_{22}	−0.058105	−16.10721	0.0000
Share Equation YL			
Adjusted R-square		0.951374	
Q-statistics (H_0: no auto-correlation)		3.2004	(p-value: 0.074)

Finance (Period: 1980–95)		1 = ICT, 2 = YL	
	Coefficient	t-statistic	Prob.
β_2	1.288469	35.75562	0.0000
γ_{22}	−0.25612	−12.6053	0.0000
Share Equation YL			
Adjusted R-square		0.956024	
Q-statistics (H_0: no auto-correlation)		0.9242	(p-value: 0.336)

Trans. & Comm. (Period: 1980–98)		1 = ICT, 2 = YL	
	Coefficient	t-statistic	Prob.
β_2	1.347934	18.51695	0.0000
γ_{22}	−0.189788	−5.36855	0.0001
1990 DUMMY for β_2	−0.031969	−2.336445	0.0338
1986 DUMMY for γ_{22}	−0.009351	−2.634962	0.0187
Share Equation YL			
Adjusted R-square		0.975456	
Q-statistics (H_0: no auto-correlation)		1.9338	(p-value: 0.164)

Services (Period: 1980–98)		1 = ICT, 2 = YL, 3 = OH	
	Coefficient	*t*-statistic	Prob.
β_2	0.488974	57.94819	0.0000
β_3	0.307996	32.08407	0.0000
γ_{22}	−0.22324	−12.31599	0.0000
γ_{23}	0.105952	11.38347	0.0000
1990 DUMMY for γ_{23}	−0.007741	−2.179989	0.0365
Share Equation YL			
Adjusted R-square		0.970672	
Q-statistics (H_0: no auto-correlation)		2.37951	(p-value: 0.123)
Share Equation OH			
Adjusted R-square		0.976691	
Q-statistics (H_0: no auto-correlation)		0.2573	(p-value: 0.612)

Notes: ICT = ICT capital. YL = young worker with low education level. YH = young worker with high education level. OH = old worker with high education level. See Table 3.1 for industry abbreviations.

Table 3.5b Estimation of cost share functions: manufacturing (six types of labor input)

Manufacturing			
Paper & pulp (Period: 1980–98)		1 = ICT, 2 = PL	
	Coefficient	*t*-statistic	Prob.
β_2	0.988995	397.2814	0.0000
γ_{22}	−0.006028	−3.157782	0.0065
1989 DUMMY for β_2	0.004776	3.021923	0.0086
1993 DUMMY for $\gamma_{22}2$	−0.003898	−4.975994	0.0002
Share Equation PL			
Adjusted R-square		0.918563	
Q-statistics (H_0: no auto-correlation)		0.0043	(p-value: 0.948)

Table 3.5b (*continued*)

Stone & clay (Period: 1980–98)		1 = ICT, 2 = PL	
	Coefficient	*t*-statistic	Prob.
β_2	1.006941	231.9819	0.0000
γ_{22}	−0.012192	−3.072057	0.0077
1985 DUMMY for β_2	−0.004007	−2.381911	0.0309
1992 DUMMY for γ_{22}	0.002725	1.93198	0.0725
Share Equation PL			
Adjusted R-square		0.875977	
Q-statistics (H_0: no auto-correlation)		0.4283	(p-value: 0.513)

Elec. machinery (Period: 1980–98)		1 = ICT, 2 = PL, 3 = NPYH	
	Coefficient	*t*-statistic	Prob.
β_2	0.62301	9.886817	0.0000
β_3	0.268787	5.332033	0.0000
γ_{22}	−0.422625	−4.451115	0.0001
γ_{23}	0.36764	3.264043	0.0027
γ_{33}	−0.297251	−2.505741	0.0177
1986 Dummy for β_2	−0.068576	−5.685736	0.0000
1989 Dummy for γ_{22}	−0.018167	−4.110175	0.0003
Share Equation PL			
Adjusted R-square		0.923523	
Q-statistics (H_0: no auto-correlation)		1.4071	(p-value: 0.236)
Share Equation NPYH			
Adjusted R-square		0.928609	
Q-statistics (H: no auto-correlation)		2.9666	(p-value: 0.085)

Trans. equipment (Period: 1990–98)		1 = IT, 2 = PL, 3 = NPOH	
	Coefficient	*t*-statistic	Prob.
β_2		0.751053	8.145365
0.0000			
β_3	0.24477	2.428891	0.0213
γ_{22}	−0.251735	−2.677556	0.0119
γ_{23}	0.212179	2.097833	0.0444
γ_{33}	−0.193118	−1.764331	0.0879

1990 DUMMY for β_2	−0.008931	−2.773712	0.0094
1988 DUMMY for β_3	−0.069378	−7.243312	0.0000
1988 DUMMY for γ_{33}	0.025617	6.569072	0.0000
Share Equation PL			
Adjusted R-square		0.941398	
Q-statistics (H_0: no auto-correlation)		0.2266	(p-value: 0.634)
Share Equation NPOH			
Adjusted R-square		0.921625	
Q-statistics (H_0: no auto-correlation)		0.0004	(p-value: 0.985)

Notes: ICT = ICT capital. PL = production worker with low education level. NPYH = nonproduction, young worker with high education level. NPOH = nonproduction, old worker with high education level.

substantially, though it still remains. Thus, the result for the electrical machinery sector is somewhat tentative and needs further investigation.

A striking result is that in all industries, old workers with low education levels are shown consistently to be quasi-fixed. Labor economists associate education level with skill level and often use education level as a proxy for skill level.[29] If this association is reasonable, then old workers with low levels of education are unskilled labor, and this part of the labor force would be expected to be variable; however, such a conclusion is not supported by our data. Rather, the result is consistent with firm-specific team-oriented skills that must be acquired over a long period: old workers with low education levels who have worked for one firm for a long time are more productive and effective on a production team than young workers of the same education levels. Because of this labor–productivity difference and externality in the workplace, the firm considers them as quasi-fixed factors rather than variable factors that can be freely adjusted to fit current economic conditions.

Table 3.4a shows that workers with high education levels are variable factors in almost half of all industries. Young well-educated workers are variable factors in the food, textile, fabricated metal and general machinery (1990s) industries, and old well-educated workers are variable factors in the instruments and services sectors. This result may seem puzzling at first sight, because Japanese firms traditionally are known for their policy of lifelong employment, and consequently one might have expected this part of the labor force to be quasi-fixed. Our result may not seem as surprising, however, if one takes account of the effect of ICT technology development.

In tables 3.6a and 3.6b, we show the results of the calculated Allen–

Table 3.6a Substitutability and complementarity. Allen–Uzawa's elasticity of substitution for baseline case (four types of labor input)

Manufacturing							
Food	1980–89	1990–98		**Textile**	1980–92	1993–98	
ICT & YL	10.445	4.5166		ICT & YL	8.7841	3.9614	
ICT & YH	−20.8430	−5.4259		ICT & YH	−22.2690	−3.7831	
YL & YH	1.4537	1.4478		YL & YH	1.5802	1.4510	
Paper & pulp	1980–88	1989–92	1993–98	**Chemicals**	1980–89	1990–98	
ICT & YL	1.6988	1.3986	1.3209	ICT & YL	3.3424	2.2828	
Stone & clay	1980–91	1992–98		**Pri. metal**	1980–87	1988–94	1995–98
ICT & YL	3.5543	1.9676		ICT & YL	3.2555	1.9228	1.6979
Fab. metal	1980–87	1988–93	1994–98	**Gen. machinery**	1980–89	1990–98	
ICT & YL	5.8058	4.1781	4.7314	ICT & YL	3.5528	6.7831	
ICT & YH	−10.695	−6.9642	−6.3244	ICT & YH		−8.6414	
YL & YH	1.3718	1.3874	1.3412	YL & YH		2.3471	
Elec. machinery	1980–92	1993–98		**Trans. equipment**	1980–88	1989–92	1993–98
ICT & YL	5.0917	3.7034		ICT & YL	2.6372	1.5929	1.4647
ICT & YH	−3.1283	−0.73924					
YL & YH	2.1956	2.1819					
Instruments	1980–93	1994–98					
ICT & YL	5.8419	4.1783					
ICT & OH	−11.169	−2.3892					
YL & OH	4.0883	3.3868					

Nonmanufacturing

Construc.	1980–89	1990–98	**Trade**	1980–98		
ICT & YL	3.7776	1.9186	ICT & YL	3.4879		
Finance	1980–95		**Trans. & commu.**	1980–85	1986–89	1990–98
ICT & YL	2.9116		ICT & YL	5.0983	3.269	2.301
Services	1980–89	1990–98				
ICT & YL	3.5142	2.1880				
ICT & OH	−1.8956	0.05846				
YL & OH	1.7061	1.1704				

Notes: ICT = ICT capital. YL = young worker with low education level. YH = young worker with high education level. OH = old worker with high education level. See Table 3.1 for industry abbreviations.

Table 3.6b Substitutability and complementarity. Allen–Uzawa elasticity of substitution for the extended case (six types of labor input)

Paper & pulp	1980–88	1989–92	1993–98
ICT & PL	1.3256	1.3347	1.3407
Stone & clay	1980–84	1985–91	1992–98
ICT & PL	2.5316	1.7619	1.5000
Elec. machinery	1980–85	1986–88	1989–98
ICT & PL	2.1819	1.6653	1.5938
ICT & NPYH	−3.4026	−1.2292	−0.1413
PL & NPYH	3.1913	3.4077	3.3558
Trans. equipment	1980–87	1988–89	1990–98
ICT & PL	4.4181	2.5665	2.2881
ICT & NPOH	−12.5610	−11.2420	−6.4093
PL & NPOH	3.5260	3.0430	2.6718

Notes: ICT = ICT capital. PL = production worker with low education level. NPYH = nonproduction, young worker with high education level. NPOH = nonproduction, old worker with high education level. See Table 3.1 for industry abbreviations.

Uzawa elasticity of substitution. In all cases in which high-education-level labor is a variable input, ICT capital stocks and high-education-level labor are complements rather than substitutes. The rapid increase in ICT investment reported in Table 3.2 induced more demand for well-educated workers capable of using ICT productively. Casual observation suggests that these well-educated workers, who specialize in information technology (for example, systems engineers) are mobile, unlike the stereotypical image of lifelong workers with high education levels. In fact, in the general machinery industry, young well-educated workers are quasi-fixed in the 1980s but become variable inputs in the 1990s. Thus, the results in Tables 3.4a and 3.4b may indicate a profound effect of information technology on the well-educated workforce, making members of that workforce variable inputs rather than quasi-fixed inputs. In contrast, but not surprisingly, young workers with high education levels are substitutes for young workers with low education levels for industries in which both are variable inputs.

Tables 3.6a and 3.6b also show differences between ICT-intensive and non-ICT-intensive industries. As explained in Section 3, the electrical machinery, instruments, finance and insurance, transportation and communication, and services industries are ICT intensive. Among them, the finance and insurance industry and the transportation and communication industry have only ICT and young workers with low levels of education as variable factors. The finance and insurance industry's data are reliable only before 1995. In the remaining three ICT-intensive industries, old well-educated workers (nonproduction old well-educated workers), rather than young well-educated workers, are complements of ICT capital stocks. We are unable to explain this difference among the ICT-intensive industries, however.

Let us now turn to the issue of dynamics with respect to the effects of advances in ICT. Tables 3.6a and 3.6b also show how substitutability and complementarity evolve in the long run. For instance, we see that the Allen–Uzawa elasticity of substitution between ICT and young workers with low levels of education decreased in the 1990s compared with the 1980s, except for in the fabricated metal sector. Similarly, the degree of complementarity between well-educated workers and ICT capital stocks also decreased. In contrast, substitutability between well-educated workers and less-educated workers did not change.

Recently, Morishima's elasticity of substitution (see Murota 1977; Kuga 1979; and Blackorby and Russell 1989), rather than Allen–Uzawa's elasticity of substitution, has been utilized in the literature (see, for example, Stiroh 1999b). We still report Allen–Uzawa's elasticity of substitution, however, because (i) Allen–Uzawa's and Morishima's are the same in two-factor cases (many industries in our sample fall into this category) and (ii) substitutability or complementarity is obvious in Allen–Uzawa's measure but not in Morishima's. We also calculated Morishima's elasticity of substitution for three-variable-factor cases. In Tables 3.7a and 3.7b, we report the results of four labor inputs case (baseline case) and six labor inputs case (manufacturing only), respectively. We found an interesting result. Morishima's elasticity shows a stark asymmetry in the effect of ICT wage-relative price changes in baseline case. If ICT prices decrease while other prices remain unchanged, then Morishima's elasticity $MES_{\mathrm{ICT,YH}}$ or $MES_{\mathrm{ICT,OH}}$ is positive, substitution. In contract, if the hourly wage of well-educated labor decreases while other prices are unchanged, then $MES_{\mathrm{YH,ICT}}$ or $MES_{\mathrm{OH,ICT}}$ is negative, complementarity. In detail, $MES_{\mathrm{ICT,YH}}$ or $MES_{\mathrm{ICT,OH}}$ of the industries except for Food (90s), Gen. machinery (90s) and Services (90s) are greater than unity, showing that if ICT prices decrease while other prices remain unchanged, the relative share of ICT increases. And in the industries (Food, Textile (80s), Fab. metal (90s) and Gen. machinery (90s)), if the hourly wage of well-educated labor decreases while other prices are unchanged, the relative share of ICT stocks increases.

Finally, let us examine the overall change of quasi-fixedness of factor inputs. In Table 3.8, the share of variable costs in the total production cost (the sum of the costs of variable factors and quasi-fixed factors) is shown

Table 3.7a Substitutability and complementarity. Morishima's elasticity of substitution for baseline case (four types of labor input)

Manufacturing

Food

1980–89	ICT	YL	YH
ICT		3.3007	2.8094
YL	8.3928		1.5949
YH	−3.9885	1.1036	

1990–98	ICT	YL	YH
ICT		1.3566	0.9147
YL	3.5413		1.5842
YH	−1.0424	1.1423	

Textile

1980–92	ICT	YL	YH
ICT		3.5562	2.9279
YL	7.5711		1.7529
YH	−2.9173	1.0977	

1993–98	ICT	YL	YH
ICT		1.9497	1.4557
YL	3.3093		1.6112
YH	−0.2424	1.1172	

Fab. metal

1980–87	ICT	YL	YH
ICT		1.9757	1.5853
YL	4.7766		1.4767
YH	−1.7147	1.0863	

1988–93	ICT	YL	YH
ICT		1.6811	1.2801
YL	3.5527		1.4878
YH	−0.7848	1.0868	

1994–98	ICT	YL	YH
ICT		1.7487	1.3894
YL	3.8168		1.4514
YH	−0.9760	1.0921	

Gen. machinery

1990–98	ICT	YL	YH
ICT		0.9903	0.5479
YL	5.1171		2.4744
YH	−2.0949	2.0319	

Elec. machinery

1980–92	ICT	YL	YH
ICT		2.8735	2.1651
YL	4.3102		2.4452
YH	0.3001	1.7368	

1993–98	ICT	YL	YH
ICT		2.0945	1.4173
YL	3.1524		2.4138
YH	0.6787	1.7367	

Instruments

1980–93	ICT	YL	OH
ICT		2.6329	1.6102
YL	5.5102		4.1928
OH	0.2928	3.1790	

1994–98	ICT	YL	OH
ICT		1.9462	1.1640
YL	3.9093		3.4810
OH	0.7358	2.6989	

Nonmanufacturing

Services

1980–89	ICT	YL	OH
ICT		1.6838	1.1110
YL	2.9024		1.8975
OH	0.1061	1.3247	

1990–98	ICT	YL	OH
ICT		1.2148	0.9023
YL	1.7229		1.3197
OH	0.4990	1.0072	

Notes: ICT = ICT capital. YL = young worker with low education level. YH = young worker with high education level. OH = old worker with high education level. See Table 3.1 for industry abbreviations. $MES_{i,j}$ is Morishima's elasticity of substitution where there is a change in price of input i p_i holding p_j constant. For example:

	ICT	YL	YH
ICT		$MES_{ICT,YL}$	$MES_{ICT,YH}$
YL	$MES_{YL,ICT}$		$MESS_{YL,YH}$
YH	$MES_{YL,ICT}$	$MES_{YH,YL}$	

Table 3.7b *Substitutability and complementarity. Morishima's elasticity of substitution for baseline case (six types of labor input)*

Manufacturing

Elec. machinery

1980–85	ICT	PL	NPYH	1986–88	ICT	PL	NPYH	1989–98	ICT	PL	NPYH
ICT		0.8502	0.4695	ICT		0.9789	0.6032	ICT		1.1168	0.8273
PL	2.4390		3.1468	PL	2.0813		3.1469	PL	2.0782		3.0618
NPYH	1.1774	2.7662		NPYH	1.6688	2.7712		NPYH	1.8109	2.7723	

Trans. equipment

1980–87	ICT	PL	NPOH	1988–89	ICT	PL	NPOH	1990–98	ICT	PL	NPOH
ICT		2.6188	2.3962	ICT		0.8747	0.4633	ICT		0.9146	0.5793
PL	4.3237		3.2630	PL	2.6248		3.0288	PL	2.3505		2.6716
NPOH	1.3355	3.0404		NPOH	0.8673	2.6175		NPOH	0.9004	2.3363	

Notes: ICT = ICT capital. PL = production worker with low education level. NPYH = nonproduction, young worker with high education level. NPOH = nonproduction, old worker with high education level. See Table 3.1 for industry abbreviations. $MES_{i,j}$ is Morishima's elasticity of substitution where there is a change in price of input i p_i holding p_j constant. For example:

	ICT	YL	YH
ICT		$MES_{ICT,YL}$	$MES_{ICT,YH}$
YL	$MES_{YL,ICT}$		$MESS_{YL,YH}$
YH	$MES_{YH,ICT}$	$MES_{YH,YL}$	

Table 3.8 Cost share of variable inputs (percentage points)

	Food	Textile	Paper & pulp	Chemicals	Stone & clay	
1981–89	0.383	0.360	0.215	0.207	0.247	
1990–98	0.329	0.265	0.180	0.162	0.193	
	Pri. metal	Fab. metal	Gen. machinery	Elec. machinery	Trans. equipment	Instruments
1981–89	0.190	0.397	0.284	0.515	0.307	0.432
1990–98	0.143	0.334	0.322	0.430	0.229	0.348
	Construc.	Trade	Finance	Trans. & commu.	Services	
1981–89	0.284	0.331	0.499	0.274	0.344	
1990–98	0.210	0.228	n.a.	0.228	0.330	

Notes: Averages in this table are arithmetic averages. Finance data in the 1990s are excluded because of data problems. See Section 3. See Table 3.1 for industry abbreviations. n.a. = not available.

for the 1980s and 1990s. The variable cost share decreased, substantially in some cases, from the 1980s to the 1990s, except for in the general machinery industry, in which the number of variable inputs increased. This is one cause of the poor performance of Japanese firms in the 1990s, when demand was very weak.

5. ICT STOCKS, HUMAN CAPITAL AND TECHNOLOGICAL PROGRESS: 1980–98

In this section, we first examine sectoral value-added growth and the contribution of each input to economic growth between 1981 and 1998. Then we derive the rate of technological progress in the framework developed in Section 3. We confirm a sharp decline in the rate of technological progress from the 1980s to the 1990s. We investigate possible causes of the decline of technological progress by examining factors that determine the rate of technological growth and analysing the ways in which they have changed over the study period. We consider 11 manufacturing industries and four nonmanufacturing industries (excluding the finance and insurance sector).[30]

As explained in Section 3, we approximate the production function parameter k by the long-run ratio of the variable cost to the total cost. In doing so, the value of k in the 1980s may be different from that in the 1990s. This is obvious for the general machinery industry, in which the number of variable inputs has changed over time. We approximate k_{1980} by the 1980s average variable cost/total cost ratio, and k_{1990} by the 1990s variable cost/total cost ratio. We then further divide the two decades into four subperiods (1981–84, 1985–89, 1990–94 and 1995–98).

Value-added Growth and Contribution of Inputs to Growth

The results on sectoral value-added growth (see Table 3.9 and Figure 3.1) reveal a remarkable contrast between the 1980s and the 1990s. Most industries show a very high rate of value-added growth in the 1980s. After the crash of the stock and real estate markets around 1990, the growth rate declines substantially and, in some industries, falls into the negative region, particularly during 1995–98. Of 11 manufacturing industries, six have the following negative rates of value-added growth during 1995–98: textile (−9.50 percent), paper and pulp (−0.64 percent), chemicals (−0.99 percent), stone and clay (−3.26 percent), primary metal (−2.43 percent) and fabricated metal (−2.30 percent). Among the four machinery industries, transportation equipment also has a negative growth rate (−0.28

percent) in that period. Three other machinery industries, however, experience a higher rate of value-added growth in the latter half of the 1990s than in the first half: general machinery (1.11 percent), electrical machinery (5.99 percent) and instruments (3.13 percent). In four nonmanufacturing industries under investigation, construction and trade have a negative rate of value-added growth during 1995–98 (−4.04 percent and −0.64 percent, respectively). The two lines in Figure 3.1 below the bar charts show the nominal GDP share of each industry.[31] The thick solid line is the nominal GDP share in 1998, and the dotted line is that in 1980. The shares of GDP of the primary metal and trade sectors declined sharply from 1980 to 1998, whereas the GDP share of services rose sharply.

The following simple regression of value-added growth on the 1990s dummy confirms a sharp decline of value-added growth from the 1980s to the 1990s. We regress the average value-added growth in four subperiods on a constant and the 1990s dummy and obtain the following result (the t-values are shown in parentheses, and N is the number of observations):

$$\textit{Value-added growth} = \underset{(7.711)}{4.773} - \underset{(-5.116)}{4.479} \times 90\text{s DUMMY},\ R^2 = 0.311,\ N = 60.$$

The 1990s dummy is very significant. A similar simple regression (not reported here) with a manufacturing dummy, however, reveals that there is no statistically significant difference between manufacturing and non-manufacturing with respect to the value-added growth pattern.

Tables 3.10 and 3.11 show each input's contribution to value-added growth. To save space, we group 'structure' and 'buildings' in the category of 'structure capital stocks', and we combine 'machines and tools' and 'transportation machines' to make the category of 'equipment capital stocks.' For labor inputs, we report the four labor input cases (young with low levels of education, young with high levels of education, old with low levels of education and old with high levels of education). Recall that the production and nonproduction classifications are not used for the non-manufacturing industries.

Table 3.10 shows that the contribution of ICT stock to value-added growth is always positive throughout the period except in the fabricated metal industry in the latter half of the 1990s, and the same is true for (non-ICT) equipment. In contrast, structure's contribution to value-added growth is small and becomes negative in the latter half of the 1990s in four industries (textile, stone and clay, primary metal, and instruments). This clearly shows that industrial growth gravitates from physical expansion to internal upgrading of equipment (both ICT-related and non-ICT-related equipment).

Table 3.11 reveals a remarkable contrast between workers with low

Table 3.9 Sources of value-added growth: 1981–98 (percentage points)

	Food	Textile	Paper & pulp	Chemicals	Stone & clay	Pri. metal	Fab. metal
Total sample period: 1981–98							
Value-added	2.288	−3.112	2.993	4.447	1.488	20.149	3.285
Variable inputs	0.163	−1.733	20.070	−0.070	−0.676	−0.209	−0.549
Quasi-fixed inputs	1.933	−0.171	2.033	1.847	0.299	1.056	0.848
Technological progress	0.088	−1.317	1.033	2.614	1.768	−0.967	3.002
1980s: 1981–89							
Value-added	2.782	−0.956	6.066	8.172	4.613	0.448	5.978
Variable inputs	0.114	−1.152	0.022	−0.435	−1.017	−0.216	−0.607
Quasi-fixed inputs	2.770	1.519	2.535	2.059	1.057	1.699	1.529
Technological progress	−0.243	−1.387	3.515	6.498	4.391	−1.033	4.975
1990s: 1990–98							
Value-added	1.796	−5.221	0.008	0.851	−1.544	−0.742	0.660
Variable inputs	0.211	−2.312	−0.161	0.296	−0.334	−0.201	−0.491
Quasi-fixed inputs	1.103	−1.833	1.534	1.636	−0.452	0.417	0.172
Technological progress	0.419	−1.248	−1.390	−1.128	−0.788	−0.902	1.066
Subperiod: 1981–84							
Value-added	3.810	−0.991	4.728	8.246	3.644	−4.403	4.606
Variable inputs	−0.018	−0.831	−0.312	−20.455	−1.290	−0.034	−1.724
Quasi-fixed inputs	2.679	1.526	2.354	1.972	0.597	2.383	0.477
Technological progress	1.052	−1.672	2.695	6.711	3.949	−6.620	5.760
Subperiod: 1985–89							
Value-added	1.967	−0.928	7.149	8.113	5.395	4.506	7.087
Variable inputs	0.220	−1.407	0.289	−0.418	−0.798	−0.362	0.295
Quasi-fixed inputs	2.843	1.513	2.680	2.128	1.427	1.154	2.379
Technological progress	−1.266	−1.158	4.175	6.328	4.745	3.676	4.351
Subperiod: 1990–94							
Value-added	2.838	−1.651	0.527	2.343	−0.148	0.625	3.090
Variable inputs	0.170	−3.174	−0.205	0.300	−0.479	−0.481	−0.382
Quasi-fixed inputs	2.770	−2.643	2.763	3.087	0.774	1.206	0.472
Technological progress	−0.134	3.978	−2.054	−1.077	−0.472	−0.114	3.068
Subperiod: 1995–98							
Value-added	0.508	−9.502	−0.636	−0.985	−3.261	−2.426	−2.298
Variable inputs	0.263	−1.223	−0.105	0.291	−0.151	0.149	−0.627
Quasi-fixed inputs	20.943	20.811	0.019	20.149	21.965	20.559	20.202
Technological progress	1.115	−7.412	−0.554	−1.192	−1.182	−1.877	−1.382

Gen. machinery	Elec. machinery	Trans. equipment	Instruments	Construc.	Trade	Finance	Trans. & commu.	Services
2.803	8.396	2.179	2.105	2.246	2.132	n.a.	2.794	4.768
−0.283	1.267	−0.369	−0.486	−0.387	−0.880	n.a.	−0.042	1.726
2.101	2.973	2.282	1.238	1.400	2.030	n.a.	2.368	3.372
1.132	4.242	0.277	1.443	1.241	0.972	n.a.	0.469	−0.348
6.437	12.152	3.969	5.018	4.300	3.463	7.917	4.682	5.462
−0.233	2.741	−0.427	−0.666	−0.667	−0.926	1.012	−0.240	1.719
3.603	4.464	3.092	2.407	1.573	2.754	1.944	3.441	3.959
3.107	4.969	1.329	3.338	3.421	1.606	4.919	1.469	−0.241
−0.706	4.765	0.420	−0.729	0.233	0.819	n.a.	0.940	4.078
−0.333	−0.186	−0.311	−0.306	−0.105	−0.835	n.a.	0.156	1.733
0.620	1.503	1.479	0.082	1.228	1.311	n.a.	1.306	2.788
−0.805	3.520	−0.765	−0.418	−0.894	0.343	n.a.	−0.521	−0.454
7.133	14.435	2.024	4.340	0.041	1.907	5.409	4.632	6.300
−0.559	4.293	−0.704	−0.888	−1.121	−0.655	1.472	−0.547	1.699
4.029	5.329	3.206	2.482	0.518	3.496	2.497	3.866	4.316
3.589	4.822	−0.475	2.758	0.640	20.936	1.418	1.292	0.284
5.884	10.359	5.551	5.565	7.837	4.725	9.966	4.722	4.796
0.028	1.516	−0.205	−0.489	−0.303	−1.142	0.646	0.006	1.734
3.264	3.777	3.000	2.347	2.425	2.164	1.504	3.103	3.675
2.723	5.086	2.795	3.805	5.702	3.687	7.806	1.610	−0.660
−2.133	3.798	0.988	−3.714	3.787	2.003	n.a.	0.878	5.119
−0.677	−0.374	−0.506	−1.023	−0.035	−0.730	n.a.	−0.047	1.703
0.732	1.910	2.064	0.395	2.146	0.846	n.a.	1.357	3.330
−1.954	2.334	−0.592	−2.942	1.632	1.900	n.a.	−0.436	0.078
1.106	5.986	−0.284	3.134	−4.038	−0.643	n.a.	1.017	2.792
0.100	0.051	−0.066	0.596	−0.193	−0.967	n.a.	0.412	1.770
0.481	0.996	0.753	20.307	0.092	1.896	n.a.	1.243	2.114
0.649	5.022	−0.980	2.829	−3.963	−1.570	n.a.	−0.626	−1.116

Notes: Averages in this table are geometric averages. The sum of input contributions and technological progress growth may not add up to value-added growth. Finance data in the 1990s are excluded because of data problems. See Section 3. See Table 3.1 for industry abbreviations. n.a. = not available.

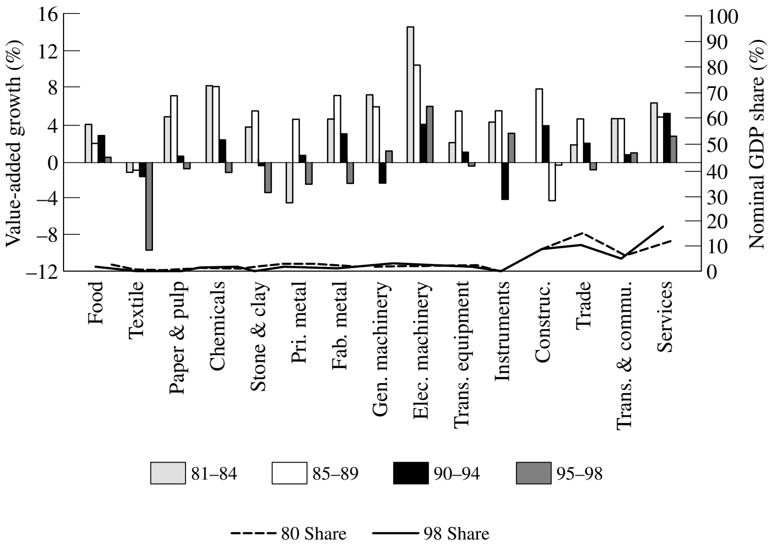


Figure 3.1 Value-added growth and GDP share

education levels and those with high education levels in the 1990s. In the 1990s, the contribution of young workers with low education levels is negative in all industries under consideration, regardless of the level of value-added growth. In contrast, the contribution of young workers with high education levels to value-added growth is positive in all industries except for the textile and instruments sector in the 1990s. Many Japanese industries are currently experiencing the effect of population ageing and are upgrading their workforce with respect to education levels. For old workers, this upgrading is far more sweeping. In the 1990s, all industries except for services have a negative contribution of old workers with low education, whereas the contribution of old workers with high education levels is positive in all industries in the same period. Thus, although old workers with low education levels are quasi-fixed (as shown in Section 4), their inputs are adjusted in the long run by natural attrition or by employment adjustment. They are quasi-fixed but variable in the long run.

Technological Progress, ICT Externality and ICT-induced Skill Obsolescence

As noted above, the rate of value-added growth declined substantially in the 1990s. This decline was not simply a result of a slump in demand and the consequent decrease in factor inputs. The rate of technological progress also declined substantially in many industries. The prolonged slump of the 1990s was not merely a demand-driven phenomenon: the supply side played a substantial role.

In Table 3.9, the rate of technological growth, which is the residual of the value-added growth that is not attributable to the inputs' contribution, is shown for the total sample period, for the 1980s and 1990s, and for four subperiods (1981–84, 1985–89, 1990–94, 1995–98). Figure 3.2 shows the changes from subperiod to subperiod, as well as the ICT stocks' share in the total capital stock of each industry in the 1981–84 period and the 1995–98 period. Table 3.9 and Figure 3.2 indicate that there is a downward shift in technological progress from the 1980s to the 1990s. To see this, we regress the subperiod average rates of technological progress on a constant and the 1990s dummy:

$$\textit{Technological progress} = \underset{(4.672)}{2.315} - \underset{(-3.733)}{2.616} \times 90\text{s DUMMY},\ R^2 = 0.194,\ N = 60. \tag{3.9}$$

The coefficient for the 1990s dummy is negative and statistically significant, suggesting a downward shift. When a dummy representing manufacturing industries is included, the coefficient of this dummy is statistically insignificant. Thus, the shift occurs in manufacturing and nonmanufacturing industries in the same way.

There are, however, a few exceptions to the general pattern of a declining rate of technological progress. The 1995–98 rate of technological progress in the electrical machinery and instruments sector is almost the same as in the 1980s. These two industries are among those industries having a high rate of ICT capital formation in both the 1980s and 1990s (Table 3.10 and Figure 3.2). This does not necessarily suggest a linkage between ICT capital formation and the rate of technological progress, however, because the services sector has a higher rate of ICT capital formation, and its rate of technological progress is negative even in the latter half of 1990s (Figure 3.2). The relationship between technological progress and ICT capital stocks is subtler, and we need to examine the issue using a more formal analysis. First, however, we review several possible factors that may influence the rate of technological progress.

First, there is a strong argument that ICT capital stocks have positive externality. Computers are interconnected via local area networks and/or the Internet. Their productivity increases more than proportionally as the number of computers increases. The value of software also increases more than proportionally as the number of users increases. Some observers argue that the US productivity increase found in Jorgenson and Stiroh (2000) Oliner and Sichel (2000), and others[32] stems partly from this externality.

Table 3.10 Capital's contribution to value-added growth (percentage points): 1981–98

	Food	Textile	Paper & pulp	Chemicals	Stone & clay	Pri. metal	Fab. metal
Total sample period: 1981–98							
ICT Capital	0.186	0.148	0.119	0.479	0.075	0.207	0.122
Equipment	0.832	0.936	1.049	0.811	0.665	0.472	0.679
Structure	0.138	−0.032	0.471	0.203	0.076	0.354	0.092
1980s: 1981–89							
ICT Capital	0.146	0.106	0.127	0.355	0.054	0.188	0.145
Equipment	1.153	0.887	0.940	0.774	0.949	0.291	0.719
Structure	0.162	−0.041	0.451	0.112	0.105	0.261	0.083
1990s: 1990–98							
ICT Capital	0.226	0.190	0.110	0.604	0.095	0.225	0.098
Equipment	0.513	0.985	1.158	0.848	0.383	0.653	0.640
Structure	0.114	−0.022	0.491	0.294	0.047	0.448	0.100
Subperiod: 1981–84							
ICT Capital	0.102	0.074	0.098	0.249	0.027	0.070	0.089
Equipment	1.075	0.507	0.591	0.685	1.026	0.005	0.443
Structure	0.130	−0.209	0.164	−0.147	0.054	−0.051	0.020
Subperiod: 1985–89							
ICT Capital	0.182	0.132	0.150	0.440	0.076	0.284	0.190
Equipment	1.215	1.192	1.221	0.845	0.887	0.521	0.939
Structure	0.188	0.093	0.682	0.320	0.145	0.511	0.134
Subperiod: 1990–94							
ICT Capital	0.255	0.173	0.138	0.537	0.064	0.229	0.181
Equipment	0.781	1.075	1.446	0.803	0.347	0.662	0.798
Structure	0.119	0.060	0.714	0.507	0.121	0.852	0.104
Subperiod 1995–98							
ICT Capital	0.190	0.211	0.076	0.688	0.133	0.221	20.006
Equipment	0.179	0.873	0.800	0.904	0.427	0.642	0.442
Structure	0.107	−0.126	0.214	0.028	−0.045	−0.055	0.095

Gen. machinery	Elec. machinery	Trans. equipment	Instruments	Construc.	Trade	Finance	Trans. & commu.	Services
0.125	0.896	0.218	0.510	0.032	0.075	n.a.	0.381	0.814
1.234	1.378	1.422	1.333	0.194	0.286	n.a.	0.967	1.371
0.102	0.488	0.300	0.045	0.093	0.268	n.a.	0.400	0.566
0.141	0.964	0.229	0.466	0.020	0.058	0.583	0.253	0.753
1.636	1.687	1.841	1.802	0.302	0.450	0.278	1.589	1.560
0.122	0.680	0.347	0.121	0.093	0.250	0.105	0.370	0.616
0.109	0.828	0.207	0.555	0.045	0.092	n.a.	0.510	0.876
0.834	1.069	1.005	0.866	0.086	0.122	n.a.	0.348	1.182
0.081	0.296	0.252	−0.032	0.092	0.285	n.a.	0.430	0.516
0.097	0.705	0.151	0.370	0.007	0.027	0.349	0.043	0.625
1.723	1.477	1.777	1.665	0.446	0.584	0.209	2.197	1.435
0.087	0.824	0.234	0.321	0.132	0.292	0.144	0.491	0.720
0.175	1.172	0.293	0.543	0.030	0.083	0.771	0.421	0.856
1.567	1.856	1.893	1.912	0.188	0.344	0.333	1.105	1.661
0.151	0.566	0.437	−0.038	0.063	0.217	0.074	0.273	0.533
0.132	0.830	0.243	0.385	0.050	0.080	n.a.	0.546	0.825
0.988	1.251	1.138	1.148	0.089	0.126	n.a.	0.273	1.526
0.117	0.394	0.455	−0.022	0.119	0.355	n.a.	0.500	0.592
0.079	0.825	0.162	0.767	0.038	0.107	n.a.	0.467	0.939
0.642	0.843	0.841	0.516	0.083	0.119	n.a.	0.443	0.755
0.036	0.173	0.000	−0.044	0.057	0.198	n.a.	0.342	0.420

Notes: Equipment = equipment capital stocks = machine and tools + transportation machines. Structure = structure capital shares = structure + buildings. Averages in this table are geometric averages. Finance data in the 1990s are excluded because of data problems. See Section 3. n.a. = not available. See Table 3.1 for industry abbreviations.

Table 3.11 Labor's contribution to value-added growth (percentage points): 1981–98

	Food	Textile	Paper & pulp	Chemicals	Stone & clay	Pri. metal	Fab. metal
Total sample period: 1981–98							
Low education, young (under 40)	−0.371	−1.777	−0.188	−0.549	−0.750	−0.416	−0.793
Low education, old (40 and over)	0.488	−1.157	0.126	0.160	−0.732	−0.047	−0.209
High education, young (under 40)	0.350	−0.094	0.162	0.237	0.073	0.090	0.125
High education, old (40 and over)	0.481	0.108	0.228	0.444	0.224	0.189	0.290
1980s: 1981–89							
Low education, young (under 40)	−0.287	−1.223	−0.106	−0.790	−1.071	−0.405	−0.906
Low education, old (40 and over)	1.091	0.502	0.724	0.350	−0.339	0.805	0.404
High education, young (under 40)	0.259	−0.033	0.189	0.202	0.077	0.076	0.157
High education, old (40 and over)	0.364	0.168	0.234	0.627	0.268	0.268	0.331
1990s: 1990–98							
Low education, young (under 40)	−0.456	−2.329	−0.271	−0.307	−0.428	−0.427	−0.680
Low education, old (40 and over)	−0.111	−2.789	−0.468	−0.031	−1.123	−0.891	−0.817
High education, young (under 40)	0.441	−0.155	0.135	0.273	0.070	0.105	0.094
High education, old (40 and over)	0.599	0.048	0.222	0.261	0.180	0.110	0.248
Subperiod: 1981–84							
Low education, young (under 40)	−0.547	−1.081	−0.410	−0.704	−1.317	−0.103	−1.783
Low education, old (40 and over)	1.193	1.157	1.212	0.501	−0.849	1.871	−0.144
High education, young (under 40)	0.427	0.178	0.174	0.225	0.043	0.192	−0.028
High education, old (40 and over)	0.278	0.068	0.215	0.714	0.335	0.368	0.164
Subperiod: 1985–89							
Low education, young (under 40)	−0.078	−1.336	0.139	−0.858	−0.874	−0.646	−0.198
Low education, old (40 and over)	1.009	−0.019	0.335	0.230	0.072	−0.040	0.844

Gen. machinery	Elec. machinery	Trans. equipment	Instruments	Construc.	Trade	Finance	Trans. & commu.	Services
−0.517	−0.117	−0.587	−1.416	−0.419	−0.955	n.a.	−0.423	−0.008
0.133	0.578	0.032	−0.185	0.089	0.228	n.a.	0.546	0.393
0.277	0.496	0.193	0.055	0.308	0.525	n.a.	0.221	1.043
0.485	0.540	0.342	0.422	0.719	0.720	n.a.	0.236	0.919
−0.374	0.882	−0.656	−1.523	−0.687	−0.984	−0.888	−0.493	0.098
0.980	1.442	0.446	0.127	0.309	0.524	0.563	1.152	0.658
0.331	0.903	0.188	0.355	0.250	0.771	1.318	0.126	1.124
0.544	0.656	0.274	0.392	0.615	0.752	0.999	0.202	0.866
−0.661	−1.105	−0.518	−1.308	−0.150	−0.927	n.a.	−0.354	−0.114
−0.706	−0.279	−0.381	−0.496	−0.131	−0.066	n.a.	−0.057	0.128
0.223	0.092	0.199	−0.244	0.366	0.279	n.a.	0.315	0.963
0.426	0.424	0.411	0.452	0.822	0.688	n.a.	0.269	0.971
−0.656	2.233	−0.854	−1.456	−1.128	−0.682	−0.457	−0.590	0.280
1.417	2.159	0.486	0.294	−0.487	0.851	1.190	0.762	0.695
0.352	1.369	0.312	0.203	0.184	1.024	1.580	0.253	1.465
0.459	0.871	0.400	0.202	0.231	0.742	0.957	0.160	0.794
−0.147	−0.186	−0.497	−1.578	−0.333	−1.225	−1.230	−0.415	−0.047
0.631	0.873	0.414	−0.007	0.951	0.263	0.064	1.465	0.629

Table 3.11 (*continued*)

	Food	Textile	Paper & pulp	Chemicals	Stone & clay	Pri. metal	Fab. metal
High education, young (under 40)	0.125	−0.203	0.202	0.184	0.103	−0.017	0.305
High education, old (40 and over)	0.433	0.249	0.250	0.557	0.215	0.188	0.466
Subperiod: 1990–94							
Low education, young (under 40)	−0.496	−2.968	−0.343	−0.236	−0.543	−0.709	−0.661
Low education, old (40 and over)	1.002	−3.419	0.059	0.392	−0.054	−0.639	−0.636
High education, young (under 40)	0.412	−0.350	0.110	0.688	0.080	0.142	0.101
High education, old (40 and over)	0.867	−0.265	0.435	0.686	0.287	0.196	0.205
Subperiod: 1995–98							
Low education, young (under 40)	−0.405	−1.523	−0.182	−0.395	−0.284	−0.073	−0.704
Low education, old (40 and over)	−1.485	−1.996	−1.122	−0.557	−2.444	−1.206	−1.044
High education, young (under 40)	0.478	0.089	0.167	−0.244	0.057	0.059	0.085
High education, old (40 and over)	0.265	0.440	−0.043	−0.269	0.046	0.002	0.301

This so-called ‘New Economy’ argument is based on the notion of externality in ICT capital stocks.[33] If there is such externality in ICT capital stocks, then the growth residual (that is, the rate of technological progress) must be correlated with ICT capital stocks in some way.

Second, casual observation shows that there is a ‘digital divide’ between the young and the old. Older workers may be skeptical about ‘new and improved’ technological gadgets and perhaps slower to adopt new technology. If such inflexibility is present in the workplace, then technological progress resulting from ICT may be lower in industries that employ more old workers than young ones.

Third, let us temporarily ignore the effect of ICT development and consider the more conventional factors that influence productivity. Skills obtained by learning by doing and on-the-job training are often considered to be the most important determinant of productivity. The so-called Toyota

Gen. machinery	Elec. machinery	Trans. equipment	Instruments	Construc.	Trade	Finance	Trans. & commu.	Services
0.314	0.531	0.089	0.476	0.304	0.569	1.109	0.025	0.852
0.613	0.485	0.172	0.544	0.923	0.759	1.032	0.236	0.925
−0.980	−1.466	−0.750	−1.667	−0.085	−0.810	n.a.	−0.592	−0.101
−0.768	−0.257	−0.185	−0.793	0.786	0.072	n.a.	−0.003	0.250
0.178	0.265	0.173	0.079	0.211	−0.033	n.a.	0.323	0.964
0.415	0.532	0.495	0.266	0.941	0.322	n.a.	0.264	0.979
−0.261	−0.652	−0.228	−0.856	−0.231	−1.073	n.a.	−0.055	−0.131
−0.627	−0.306	−0.625	−0.124	−1.267	−0.239	n.a.	−0.125	−0.025
0.280	−0.124	0.231	−0.647	0.559	0.670	n.a.	0.304	0.962
0.440	0.288	0.305	0.684	0.674	1.147	n.a.	0.276	0.961

Notes: Averages in this table are geometric averages. Finance data in the 1990s are excluded because of data problems. See Section 3. See Table 3.1 for industry abbreviations.
n.a. = not available.

production system, which combines *kanban* (just-in-time) and TQC (total quality circle), clearly recognizes the importance of such skills acquired on site. Long-term knowledge about jobs and coworkers greatly enhances a worker's productivity in team production. This is externality in the workplace, and one worker's productivity is positively related to his/her coworkers' productivity. If this is important in production, industries that employ numerous old workers with many years of experience would be expected to show a higher growth residual (technological progress). This productivity advantage might be eroded by advances in information technology, however, if such innovations render this tacit knowledge obsolete. Thus, if this mechanism is important, one would expect a positive correlation between the ratio of old workers and the growth residual before the rapid increase of information capital stocks and a negative correlation after it.

Fourth, a classical Schumpeterian argument states that technological

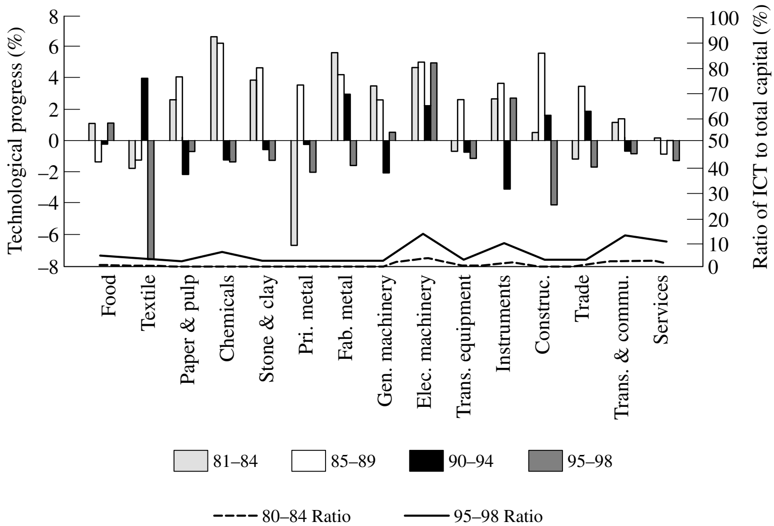


Figure 3.2 Technological progress and ICT ratio

development is often carried out by monopolistic firms. If this is the case, one would expect a positive correlation between pure profits and the growth residual. An alternative argument is that monopolistic firms generally do not experience any market pressure to innovate, so one would expect a negative correlation between pure profits and the growth residual. Finally, there is a strong argument that the impact of capital stocks on structure is different from its impact on equipment.[34] We shall also consider this possibility.

To examine the validity of the above arguments, we employ the panel data of the 15 industries and four subperiods used in our previous analysis. We then estimate an equation explaining the growth residual, or equivalently, the rate of technological progress, by (i) the ratio of old workers with low levels of education to the total labor inputs (OL), (ii) the ratio of old workers with high levels of education to the total labor inputs (OH), (iii) the ratio of a net profit to the total cost (PROFIT), (iv) the ratio of ICT stocks to the total capital stocks (ITK), and (v) the ratio of the non-ICT equipment capital stocks to total capital stocks (EQ), as follows:[35]

$$\begin{aligned} \textit{Technological progress} = \textit{constant} &+ (\beta_{\mathrm{OL}} + \delta_{\mathrm{OL}} \times \text{90s DUMMY}) \times \mathrm{OL} \\ &+ (\beta_{\mathrm{OH}} + \delta_{\mathrm{OH}} \times \text{90s DUMMY}) \times \mathrm{OH} \\ &+ (\beta_{\mathrm{PROFIT}} \times \textit{PROFIT}) + (\beta_{\mathrm{ITK}} \times \textit{ITK}) + (\beta_{\mathrm{EQ}} \times \textit{EQ}) + \varepsilon_{it}. \end{aligned}$$

Here we allow the possibility of structural change resulting from ICT development around 1990 by including a coefficient dummy variable for the 1990s (90s DUMMY). Since we are not sure about stochastic properties of industry-specific effects, we estimate both the fixed effect model and the random effect model. In addition, since explanatory variables may be endogenously determined so that they may be correlated with error terms, we also employ the generalized method of moments (GMM).[36]

Our results are shown in Tables 3.12–15. In all cases the random effect model is chosen by the Hausman test, so we report only the random effect model here. The use of GMM reveals that the coefficient of 90s DUMMY$\times$OL, δ_{OL} is not statistically significant and to include this variable makes other estimates deteriorate, so we exclude it from the GMM part of Tables 3.11–14. Similar problems occur for PROFIT, so it is also excluded from the GMM part of Tables 3.11–14.

Table 3.12 reports the results for the case in which we use only manufacturing industries. The number of observations (N) is 44. The table shows that $\beta_{OH}>0$, $\delta_{OH}<0$ and $\beta_{ITK}>0$ are statistically significant at the 10 percent, 1 percent and 5 percent levels, respectively. Thus, the result supports the existence of positive ICT externality (New Economy effect), a positive effect on productivity growth of old workers with many years of experience in the high-education segment in the 1980s (long-term employment effect), and a negative effect of ICT in the 1990s on the productivity growth of manufacturing firms with old workers with many years of experience (ICT-induced obsolescence effect). In contrast, we find no effect on manufacturing firms' productivity growth that can be attributed to the inflexibility of older workers, no pure-profit effect, and no evidence of externality in (non-ICT) equipment. Qualitatively, the same result is obtained for all industries, reported in Table 3.15. It should be noted, however, that when the sample industries are four nonmanufacturing industries, no variable has explanatory power with respect to the rate of technological progress (Table 3.14).

However, the positive ICT externality (New Economy) effect is not robust. Let us exclude electrical machinery from these 11 manufacturing industries and restrict sample industries to 10 manufacturing industries ($N=40$). The result is reported in Table 3.13. The coefficient (β_{ITK}) is now statistically insignificant. In contrast, the obsolescence effect of ICT (δ_{OH}) is still statistically significant. Thus, ICT's effects are mostly concentrated in the electrical machinery industry, which is an ICT-producing industry, and there is no compelling evidence for *general* ICT externality. It should be kept in mind, however, that the ICT externality we examined is a within-the-industry externality: for example, a network effect in the same industry. A cross-industry spillover of ICT effects may occur: for example, one

Table 3.12 Technological growth, old workers and ICT in manufacturing industries

Parameter	Estimate	Standard error	*t*-statistic	*p*-value
Constant				
Random effects	0.349	4.957	0.070	0.944
GMM	0.860	4.141	0.208	0.836
β_{OL}				
Random effects	0.011	0.081	0.137	0.891
GMM	0.018	0.051	0.356	0.722
δ_{OL}				
Random effects	0.015	0.036	0.426	0.670
GMM	n.a.	n.a.	n.a.	n.a.
β_{OH}				
Random effects	0.644	0.370	1.740	0.082
GMM	0.660	0.430	1.535	0.125
δ_{OH}				
Random effects	20.861	0.287	22.995	0.003
GMM	20.782	0.250	23.125	0.002
β_{PROFIT}				
Random effects	1.541	2.918	0.528	0.597
GMM	n.a.	n.a.	n.a.	n.a.
δ_{ITK}				
Random effects	0.515	0.226	2.282	0.022
GMM	0.392	0.153	2.568	0.010
β_{EQ}				
Random effects	20.045	0.056	20.808	0.419
GMM	20.046	0.036	21.281	0.200

Specification test	Value	*p*-value
Hausman (FE vs. RE)	10.353	0.169
Hansen	1.493	0.684

Notes: Number of industries = 11. Number of observations = 44. Dependent variable = rate of technological progress. n.a. = not applicable. FE = fixed effect. RE = random effect. Hansen = Hansen's overidentifying restrictions test (Hansen 1982).

Table 3.13 Technological growth, old workers and ICT in manufacturing industries, excluding electrical machinery

Parameter	Estimate	Standard error	t-statistic	p-value
Constant				
Random effects	1.151	6.363	0.181	0.856
GMM	0.291	5.174	0.056	0.955
β_{OL}				
Random effects	0.002	0.098	0.023	0.982
GMM	0.020	0.060	0.328	0.743
δ_{OL}				
Random effects	0.022	0.041	0.526	0.599
GMM	n.a.	n.a.	n.a.	n.a.
β_{OH}				
Random effects	0.570	0.422	1.351	0.177
GMM	0.673	0.455	1.481	0.139
δ_{OH}				
Random effects	−0.898	0.313	−2.874	0.004
GMM	−0.745	0.279	−2.675	0.007
β_{PROFIT}				
Random effects	2.147	3.259	0.659	0.510
GMM	n.a.	n.a.	n.a.	n.a.
β_{ITK}				
Random effects	0.684	0.460	1.487	0.137
GMM	0.146	0.404	0.361	0.718
β_{EQ}				
Random effects	−0.053	0.066	−0.806	0.421
GMM	−0.030	0.041	−0.744	0.457
Specification test	**Value**		**p-value**	
Hausman (FE vs. RE)	9.279		0.233	
Hansen	1.815		0.612	

Notes: Number of industries = 10. Number of observations = 40. Dependent variable = rate of technological progress. n.a. = not applicable. FE = fixed effect. RE = random effect. Hansen = Hansen's overidentifying restrictions test (Hansen 1982).

Table 3.14 Technological growth, old workers, and ICT in nonmanufacturing industries

Parameter	Estimate	Standard error	*t*-statistic	*p*-value
Constant				
Random effects	−0.185	6.961	−0.027	0.979
GMM	−1.868	4.949	−0.377	0.706
β_{OL}				
Random effects	0.083	0.212	0.392	0.695
GMM	0.089	0.138	0.644	0.520
δ_{OL}				
Random effects	−0.073	0.084	−0.871	0.383
GMM	n.a.	n.a.	n.a.	n.a.
β_{OH}				
Random effects	−0.144	0.719	−0.201	0.841
GMM	0.184	0.624	0.295	0.768
δ_{OH}				
Random effects	0.043	0.368	0.116	0.908
GMM	−0.082	0.251	−0.326	0.744
β_{PROFIT}				
Random effects	7.279	8.078	0.901	0.368
GMM	n.a.	n.a.	n.a.	n.a.
β_{ITK}				
Random effects	0.144	0.359	0.401	0.689
GMM	−0.295	0.194	−1.522	0.128
β_{EQ}				
Random effects	20.050	0.148	20.339	0.735
GMM	20.014	0.086	20.166	0.868
Specification test	**Value**	***p*-value**		
Hausman (FE vs. RE)	3.975	0.409		
Hansen	5.534		0.137	

Notes: Number of industries = 4. Number of observations. = 16. Dependent variable = rate of technological progress. FE = fixed effect. RE = random effect. Hansen = Hansen's overidentifying restrictions test (Hansen 1982). n.a. = not applicable.

Table 3.15 Technological growth, old workers, and ICT in all industries

Parameter	Estimate	Standard error	*t*-statistic	*p*-value
Constant				
Random effects	0.514	3.285	0.157	0.876
GMM	1.253	3.914	0.320	0.749
β_{OL}				
Random effects	−0.016	0.058	−0.275	0.783
GMM	−0.007	0.052	−0.143	0.886
δ_{OL}				
Random effects	−0.009	0.031	−0.277	0.781
GMM	n.a.	n.a.	n.a.	n.a.
β_{OH}				
Random effects	0.116	0.287	0.402	0.688
GMM	0.514	0.433	1.188	0.235
δ_{OH}				
Random effects	−0.399	0.218	−1.826	0.068
GMM	−0.556	0.229	−2.424	0.015
β_{PROFIT}				
Random Effects	4.099	2.158	1.899	0.058
GMM	n.a.	n.a.	n.a.	n.a.
β_{ITK}				
Random effects	0.280	0.140	1.991	0.046
GMM	0.031	0.146	0.212	0.832
β_{EQ}				
Random effects	0.011	0.036	0.292	0.770
GMM	−0.018	0.030	−0.610	0.542
Specification test	**Value**	***p*-value**		
Hausman (FE vs. RE)	6.536	0.479		
Hansen	2.653	0.448		

Notes: Number of industries = 15. Number of observations = 60. Dependent variable = rate of technological progress. n.a. denotes not applicable. FE = fixed effect. RE = random effect. Hansen = Hansen's overidentifying restrictions test (Hansen 1982).

industry's ICT investment may represent a positive externality for other industries. We do not have a clear model of technology externality over industry borders and must leave such investigation to future research.

This result is consistent with the conclusion of Stiroh (2001), who uses data for US manufacturing industries from 1973 to 1999 to estimate the correlation between ICT capital intensity and the rate of technological progress. His results suggest that the primary impact of ICT is through traditional capital-deepening and provide little evidence that ICT capital formation is responsible for accelerating the rate of technological progress in the United States.

Let us summarize the results obtained in this section. For the effect of labor force composition on the rate of technological progress, the results do not support the claim that inflexible old workers played a role in Japan's productivity slowdown. There is no correlation between the rate of technological progress and the ratio of old workers with low levels of education in the total labor inputs. The results suggest, however, that ICT development in the 1990s had a negative impact on the past strength of the Japanese economy. This past strength may be attributed to productivity increases gained from well-educated workers' learning by doing. In the manufacturing industries that have experienced strong growth in the past, the rate of technological progress in the 1980s has a positive (though weak) correlation with a 'maturing' well-educated labor force. That is, the ratio of older well-educated workers in the total labor inputs has a positive (though weak) effect on technological progress. This suggests that the increased average skill among well-educated workers resulting from longer experience has a positive effect on productivity. The relationship changes significantly in the 1990s, however, and becomes negative. The nature of technological progress apparently changed adversely. We find no evidence to support ICT externality, except in the case of the ICT-producing industries.

6. CONCLUDING REMARKS ON POLICY IMPLICATIONS

The improved data set we compiled reveals that ICT capital stocks are an important substitute for young workers with low education levels (Section 4). These results strongly suggest that ICT investment is an effective way to counter the prospective shortage of young workers in Japan. The results also imply that to strengthen this effect of ICT investment, it is necessary to improve the educational level of the labor force; otherwise, the impact of ICT investment may be seriously hindered by a shortage of complementary well-educated labor inputs. The need to improve education levels is all the more apparent if one considers the fact that ICT capital stocks and high-education-level labor are complements in all cases in which high-education-level labor is a variable input.

The results of Section 5, however, show that the hope that many economists and politicians have with respect to the 'ICT revolution', in which ICT externality greatly enhances productivity, is not supported by the data, at least with respect to within-industry effects. The productivity gain in ICT-producing industries is remarkable (for example, the electrical machinery industry in our sample), but this is an industry-specific phenomenon rather than a revolution that affects all industries. On the contrary, our results suggest that IT has a negative indirect effect on productivity. The advent of information technology may have drastically changed the comparative advantages of Japan's industries. The past technological and managerial strengths of Japanese manufacturing firms, which have been based on workers' learning by doing in the workplace and other strategies (such as TQC and on-the-job/off-the-job training), may no longer be advantages as knowledge management systems improve and become easily transferred across international borders.

In this respect, Japan needs a thorough examination of its productivity slowdown in the 1990s, especially of the strengths and weaknesses in technology and management. As our data suggest, technology and management are not independent of one another. In Section 5, we have shown that the long-term employment relationship that traditional Japanese management has cherished for the past 40 years increased productivity growth until the advent of the ICT revolution of the 1990s. This means that Japanese management was best suited for the technology available at that time, that is, pre-ICT production technology. After the ICT revolution, however, we observe the long-term employment relationship (and thus this part of Japanese management) becoming a stumbling block for productivity growth. This suggests that one form of management (including work organization and personnel management) may be efficient for one form of technology but not for others. Management styles are often stable in the long run, and mismatches may evolve between management and current technology.[37]

Moreover, technology itself is not exogenous. The past history of technological development shows the importance of the government in enhancing particular types of technological development. Obviously the government cannot choose appropriate technology for the economy, but it can provide a menu of possible technological choices from which the market chooses the winning one. The government can influence the course of technological development and properties of production technology, though it cannot determine them.

However, as the government delineates the choices and exerts its influence, *appetence* between technology and management should be properly taken into consideration. As we explained above and in Section 5, a long-term relationship is a particularly effective management strategy for technology in which tacit knowledge acquired from long experience plays a crucial role in efficiently producing products of 'integral architecture,' such as cars.[38] Simply assembling a large number of car parts does not guarantee a good car. We need tacit knowledge of design and production to make these parts work properly and smoothly long after production. A good Toyota car is thus an 'integral product', far more than just assembled 'modular parts'. However, this is not the case for personal computers. (One can assemble a very good desktop computer from a central processing unit, random-access memory modules, hard disk drives and so on.) We do not need tacit knowledge acquired from long experience to assemble computer parts: even with little experience one can assemble all components properly and produce a good personal computer if one has a good guidebook or guide video. In the ICT industries this modular architecture is often coupled with very rapid changes in product specification of modular parts. In such industries, a long-term relationship is likely to be a disadvantage rather than an advantage.

The advent of information and communication technology and competitive modular-architecture products marked the end of the happy coupling of traditional Japanese management with integral-architecture technology. The challenge that the Japanese economy faces is thus not only to adjust to modular-product technology, but also to find new integral-architecture technology in which Japanese management has a competitive edge.

APPENDIX 3A1 DERIVATION OF MULTIPLICATIVELY SEPARABLE VARIABLE COST FUNCTION

From Assumption 2, we have:

$$Y = G(x_1, \ldots, x_i, \ldots, x_n; A)S = G\left(1, \frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}; A\right) x_1^k S.$$

Consequently, we obtain:

$$C_V(p_1, \ldots, p_n, Y, S; A = \underset{x_1, \ldots, x_n}{\text{Min}} \left(1 + \sum_{i=2}^{n} \frac{p_i x_i}{p_1 x_1}\right) p_1 x_1$$

$$\text{subject to } \frac{Y}{x_1^k S} = G\left(1, \frac{x_2}{x_1}, \ldots, \frac{x_n}{x_1}; A\right).$$

Then, the cost minimization has three steps. In the first step, for given x_1 and Y, the ratios $\{x_i/x_1\}$ are optimized. Let $v_i{}^*$ be the resulting optimum ratio, such that:

$$\left(\frac{x_i}{x_1}\right)^* = v_i^*\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}, \frac{Y}{x_1^k S}; A\right) \text{ for } i=2\ldots, n.$$

In the second step, the optimal x_1^* is implicitly determined by:

$$\frac{Y}{x_1^k S} = G\left[1, v_2^*\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}, \frac{Y}{x_1^k S}; A\right), \ldots, v_n^*\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}, \frac{Y}{x_1^k S}; A\right)\right].$$

Finally, the optimal x_i^* is determined by $x_i^* = v_i^* x_1^*$.

Let us now show that the variable cost function C_v is multiplicatively separable between relative prices on the one hand, and output and production capacity on the other. Let h such that:

$$h = h\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}; A\right)$$

be the solution of:

$$h = G\left[1, v_2^*\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}, h; A\right) \ldots, v_n^*\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}, h; A\right)\right].$$

Note that h is a function of only the relative variable input prices (and the state of production technology A). Then we have $Y/(x_1{}^k S) = h$, which in turn implies:

$$x_1 = \left(\frac{Y}{hS}\right)^{1/k}.$$

Substituting these results into the variable cost function, we have:

$$C_V(p_1, \ldots, p_n, Y, S; A) = \left[1 + \sum_{i=2}^{n} \frac{p_i}{p_1} \tilde{v}_i^*\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}; A\right)\right] \frac{p_1}{h^{1/k}}\left(\frac{Y}{S}\right)^{1/k}$$

where:

$$\tilde{v}_i^*\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}; A\right) = v_i^*\left[\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}, h\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}; A\right); A\right].$$

Consequently, under Assumptions 1 and 2, we have multiplicatively separable variable cost function (3.3) such that:

$$C_V(p_1, \ldots, p_n, Y, S; A) = c_v(p_1, \ldots, p_n; A)\left(\frac{Y}{S}\right)^{1/k}$$

where c_v is homogeneous of degree one in prices such that:

$$c_v(p_1, \ldots, p_n; A)$$

$$= \left[1 + \sum_{i=2}^{n} \frac{p_i}{p_1} \tilde{v}_i^* \left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}; A \right) \right] p_1 \left[h\left(\frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}; A \right) \right]^{-1/k} . \quad (3A1.1)$$

APPENDIX 3A2 MULTIPLICATIVELY SEPARABLE TRANSLOG VARIABLE COST FUNCTION, SHARE EQUATIONS AND ELASTICITY OF SUBSTITUTION

We are concerned with the following form of n-factor multiplicatively-separable variable cost functions.

$$C_V(p_1, \ldots, p_n, Y, S; A) = c_v(p_1, \ldots, p_n; A) \left(\frac{Y}{S} \right)^{1/k} . \quad (3A2.1)$$

We assume that a translog approximation of c_v at $\bar{p} = (\bar{p}_1, \ldots, \bar{p}_n)$ is a good approximation. In many applications of translog functions and in many textbooks, prices are normalized through appropriate choice of units, either by setting a particular year's price equal to unity or by making the average price equal to unity, and then set $\bar{p} = (1, \ldots, 1)$. This makes exposition simple and straightforward in the traditional share equation estimation.

In this chapter, however, we do not normalize prices and we let $\bar{p}$ be the average price vector. Thus, we have $\bar{p} \neq (1, \ldots, 1)$ in general in this chapter. We adopt this procedure since parameter estimations are not invariant with respect to normalization. We get sharper results without normalization than with normalization.

Taking logarithm of c_v and then taking a second-order Taylor expansion of $1\text{n}\, c_v$ with respect to $1\text{n}\, p_i$ around $p = \bar{p}$, we have:

$$1\text{n}\, c_v(p_1, \ldots, p_n; A) = \alpha(A) + \sum_{i=1}^{n} \beta_i(A) 1\text{n}\, p_i + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \gamma_{ij}(A) 1\text{n} p_i 1\text{n}\, p_j$$

where:

$$\alpha(A) = 1\text{n}\, c_v(\bar{p}_1, \ldots, \bar{p}_n) - \sum_{i=1}^{n} \left. \frac{\partial \ln c_v}{\partial \ln p_i} \right|_{p=\bar{p}} (1\text{n}\bar{p}_i) +$$

$$\frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \left. \frac{\partial^2 \ln c_v}{\partial \ln p_i \, \partial \ln p_j} \right|_{p=\bar{p}} (1\text{n}\bar{p}_u)(1\text{n}\bar{p}_j);$$

$$\beta_i(A) = \frac{\partial \ln c_v}{\partial \ln p_i}\bigg|_{p=\bar{p}} - \sum_{j=1}^{n} \frac{\partial^2 \ln c_v}{\partial \ln p_i \, \partial \ln p_j}\bigg|_{p=\bar{p}} (1\mathrm{n}\bar{p}_j);$$

and

$$\gamma_{ij}(A) = \frac{\partial^2 \ln c_v}{\partial \ln p_i \, \partial \ln p_j}\bigg|_{p=\bar{p}}.$$

Note that all these parameters depend on the state of production technology A.

Let us examine the requirement that the cost function is homogeneous of degree one in input prices. Since C_v has the form (3A2.1), it is obvious that c_v should be homonegeneous of degree one in input prices. As usual, this implies:

$$\begin{aligned}
\ln c_v(\lambda p_1, \ldots, \lambda p_n; A) &= \alpha(A) + \sum_{i=1}^{n} \beta_i(A)(\ln\lambda + \ln p_i) \\
&\quad + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}(A)(\ln\lambda + \ln p_i)(\ln\lambda + \ln p_j) \\
&= \alpha(A) + \left(\sum_{i=1}^{n} \beta_i(A)\right)\ln\lambda + \sum_{i=1}^{n} \beta_i(A)\ln p_i \\
&\quad + \frac{1}{2}\sum_{j=1}^{n}\left[\left(\sum_{i=1}^{n} \gamma_{ij}(A)\right)\ln\lambda\right]\ln p_j \\
&\quad + \frac{1}{2}\sum_{i=1}^{n}(\ln p_i)\left[\left(\sum_{i=1}^{n} \gamma_{ij}(A)\right)\ln\lambda\right] \\
&\quad + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}(A)\ln p_i \ln p_j \\
&= \ln \lambda c_v(p_1, \ldots, p_n; A) = \ln\lambda + \alpha(A) \\
&\quad + \sum_{i=1}^{n} \beta_i(A)\ln p_i + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} \gamma_{ij}(A)\ln p_i \ln p_j.
\end{aligned}$$

As in the standard case, we have the following requirement:

$$\sum_{i=1}^{n} \beta_i(A) = 1, \ \sum_{i=1}^{n} \gamma_{ij}(A) = 0, \ \sum_{j=1}^{n} \gamma_{ij}(A) = 0, \ \gamma_{ij}(A) = \gamma_{ij}(A), \text{ for all } i \text{ and } j. \tag{3A2.2}$$

Regression Equations: Share Functions

Although our formulation of cost function deviates slightly from the standard one, we have the same form of share functions. To see this, note that:

$$\ln c_v(p_1, \ldots, p_n; A) = \ln\left[p_1 \cdot c_v\left(1, \frac{p_2}{p_1}, \ldots, \frac{p_n}{p_1}; A\right)\right]$$

$$= \ln p_1 + \left[\sum_{i=2}^{n} \beta_i(A) \ln\frac{p_i}{p_1} + \frac{1}{2}\sum_{i=2}^{n}\sum_{j=2}^{n} \gamma_{ij}(A) \ln\frac{p_i}{p_1} \ln\frac{p_j}{p_1}\right]$$

$$= \ln p_1 + \sum_{i=2}^{n} \beta_i(A)(\ln p_i - \ln p_1)$$

$$+ \frac{1}{2}\sum_{i=2}^{n}\sum_{j=2}^{n} \gamma_{ij}(A)(\ln p_i - \ln p_1)(\ln p_j - \ln p_1).$$

Let x_i be the quantity of the ith input. Then, we get:

$$\frac{\partial C_V}{\partial p_i} = x_i = \left(\frac{\partial c_v}{\partial p_i}\right)\left(\frac{Y}{S}\right)^{1/k}.$$

Combining these two relations, we get the following representation of share functions, for $i = 2, \ldots, n$:

$$s_i = \frac{p_i x_i}{C_V} = \frac{p_i x_i}{c_v(p_1, \ldots, p_n)\left(\frac{Y}{S}\right)^{1/k}} = \frac{p_i}{c_v(p_1, \ldots, p_n)}\left(\frac{\partial c_v}{\partial p_i}\right)$$

$$= \frac{d \ln c_v}{d \ln p_i} = \beta_i(A) + \sum_{j=2}^{n} \gamma_{ij}(A)(\ln p_i - \ln p_1). \qquad (3A2.3)$$

These equations are regression equations from which the following parameters of the cost function can be retrieved:

$$(\beta_i(A), \gamma_{ij}(A) \mid i, j = 2, \ldots, n).$$

From these estimated $\beta_i(A)$ and $\gamma_{ij}(A)$, the rest of parameters are calculated using the homogeneity relations (3A2.2).

First- and Second-order Derivatives

To avoid heavy notations, we hereafter suppress '(A)' in α, β_i, γ_{ij}.

First-order Derivate
It is immediate to have:

$$\frac{\partial c_v}{\partial p_i} = \left(\frac{c_v}{p_i}\right)\left(\frac{\partial \ln c_v}{\partial \ln p_i}\right) = \left(\frac{c_v}{p_i}\right)\left(\beta_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j\right). \tag{3A2.4}$$

Second-order derivative
Note that:

$$\frac{\partial \ln c_v}{\partial \ln p_i} = \beta_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j.$$

Thus we have:

$$\frac{\partial}{\partial \ln p_i}\left(\frac{\partial \ln c_v}{\partial \ln p_i}\right) = \gamma_{ii} = p_i \frac{\partial}{\partial p_i}\left(\frac{p_i}{c_v}\frac{\partial c_v}{\partial p_i}\right)$$

$$= p_i\left\{\left[\frac{\partial}{\partial p_i}\left(\frac{p_i}{c_v}\right)\right]\frac{\partial c_v}{\partial p_i} + \frac{p_i}{c_v}\frac{\partial}{\partial p_i}\left(\frac{\partial c_v}{\partial p_i}\right)\right\}$$

$$= p_i\left\{\left[\left(\frac{c_v - p_i \dfrac{\partial c_v}{\partial p_i}}{c_v^2}\right)\right]\frac{\partial c_v}{\partial p_i} + \frac{p_i}{c_v}\left(\frac{\partial^2 c_v}{\partial p_i^2}\right)\right\}$$

$$= \left(1 - \frac{p_i}{c_v}\frac{\partial c_v}{\partial p_i}\right)\frac{p_i}{c_v}\frac{\partial c_v}{\partial p_i} + \frac{p_i^2}{c_v}\left(\frac{\partial^2 c_v}{\partial p_i^2}\right)$$

which implies

$$\frac{\partial^2 c_v}{\partial p_i^2} = \frac{c_v}{p_i^2}\left[\gamma_{ij} + \left(\frac{\partial \ln c_v}{\partial \ln p_i} - 1\right)\frac{\partial \ln c_v}{\partial \ln p_i}\right]. \tag{3A2.5}$$

Similarly, we have:

$$\frac{\partial}{\partial \ln p_j}\left(\frac{\partial \ln c_v}{\partial \ln p_i}\right) = \gamma_{ij} = p_j \frac{\partial}{\partial p_j}\left(\frac{p_i}{c_v}\frac{\partial c_v}{\partial p_i}\right)$$

$$= p_j\left\{\left[\frac{\partial}{\partial p_j}\left(\frac{p_i}{c_v}\right)\right]\frac{\partial c_v}{\partial p_i} + \frac{p_i}{c_v}\frac{\partial}{\partial p_j}\left(\frac{\partial c_v}{\partial p_i}\right)\right\}$$

$$= p_j\left\{\left[p_i\left(-\frac{1}{c_v^2}\frac{\partial c_v}{\partial p_j}\right)\right]\frac{\partial c_v}{\partial p_i} + \frac{p_i}{c_v}\left(\frac{\partial^2 c_v}{\partial p_i \partial p_j}\right)\right\}$$

$$= -\left(\frac{p_j}{c_v}\frac{\partial c_v}{\partial p_j}\right)\frac{p_i}{c_v}\frac{\partial c_v}{\partial p_i} + \frac{p_i p_j}{c_v}\partial\frac{\partial^2 c_v}{\partial p_i \partial p_j}$$

which implies:

$$\frac{\partial^2 c_v}{\partial p_i \partial p_j} = \frac{c_v}{p_i p_j}\left[\gamma_{ij} + \left(\frac{\partial \ln c_v}{\partial \ln p_j}\right)\frac{\partial \ln c_v}{\partial \ln p_i}\right]. \tag{3A2.6}$$

The above discussion reveals the following simple characterization of the second derivatives. Let n_i be the input-price elasticity of the cost function:

$$\eta_i = \frac{\partial \ln c_v}{\partial \ln p_i} = \beta_i + \sum_{j=1}^{n}\gamma_{ij}\ln p_j.$$

As has been shown in the translog case, the input-price elasticity happens to be equal to the cost share of the inputs:

$$\eta_i = \frac{\partial \ln c_v}{\partial \ln p_i} = \frac{p_i x_i}{C_V} = s_i.$$

Then we have:

$$z_{ii} = \frac{\partial^2 c_v}{\partial p_i^2} = \frac{c_v}{p_i^2}\tilde{z}_{ii} \tag{3A2.7}$$

where:

$$\tilde{z}_{ii} = \gamma_{ii} + (\eta_i - 1)\eta_i \tag{3A2.8}$$

and

$$z_{ij} = \frac{\partial^2 c_v}{\partial p_i \partial p_j} = \frac{c_v}{p_i p_j}\tilde{z}_{ij} \quad i \neq j \qquad \text{(3A2.9}$$

where:

$$\tilde{z}_{ij} = \gamma_{ij} + \eta_i \eta_j. \tag{3A2.10}$$

Cost-function Requirements

By construction, homogeneity of degree one with respect to input prices is satisfied. The remaining requirements are monotonicity and concavity.

Monotonicity

The monotonicity requirement is satisfied if:

$$\frac{\partial c_v}{\partial p_i} > 0$$

Since $c_v > 0$ and $p_i > 0$, we have from (3A2.4):

$$\operatorname{sgn}\left(\frac{\partial c_v}{\partial p_i}\right) = \operatorname{sgn}\left(\frac{\partial \ln c_v}{\partial \ln p_i}\right) = \operatorname{sgn}\left(\frac{p_i x_i}{C_V}\right) = \operatorname{sgn}\left(\beta_i + \sum_{j=1}^{n} \gamma_{ij} \ln p_j\right).$$

Consequently, if the share is positive for the range of variables we observe, then monotonicity is satisfied.

Concavity

The concavity requirement is satisfied if for all *i* and *j*, the following relations are satisfied in the case of four variables,

$$z_{ii} < 0 \tag{3A2.11}$$

$$\det \begin{bmatrix} z_{ii} & z_{ij} \\ z_{ji} & z_{jj} \end{bmatrix} > 0; \tag{3A2.12}$$

$$\det \begin{bmatrix} z_{ii} & z_{ij} & z_{ik} \\ z_{ji} & z_{jj} & z_{jk} \\ z_{ki} & z_{kj} & z_{kk} \end{bmatrix} < 0. \tag{3A2.13}$$

The concavity requirement in the general case of *n* is analogously derived. Since *n* is at most four in this chapter, we examine the four-factor case here.

Note that since $c_v > 0$and $p_i > 0$, we have:

$$\operatorname{sgn}(z_{ii}) = \operatorname{sgn}\left(\frac{c_v}{p_i^2}\tilde{z}_{ii}\right) = \operatorname{sgn}(\tilde{z}_{ii})$$

$$\begin{aligned} \operatorname{sgn}\left(\det \begin{bmatrix} z_{ii} & z_{ij} \\ z_{ji} & z_{jj} \end{bmatrix}\right) &= \operatorname{sgn}[z_{ii} z_{jj} - z_{ij}^2] = \operatorname{sgn}\left[\frac{c_v}{p_i^2}\tilde{z}_{ii}\frac{c_v}{p_j^2}\tilde{z}_{jj} - \left(\frac{c_v}{p_i p_j}\right)^2 (\tilde{z}_{ij})^2\right] \\ &= \operatorname{sgn}\left\{\left(\frac{c_v}{p_i p_j}\right)^2 [\tilde{z}_{ii}\tilde{z}_{jj} - (\tilde{z}_{ij})^2]\right\} = \operatorname{sgn}\{[\tilde{z}_{ii}\tilde{z}_{jj} - (\tilde{z}_{ij})^2]\} \\ &= \operatorname{sgn}\left(\det \begin{bmatrix} \tilde{z}_{ii} & \tilde{z}_{ij} \\ \tilde{z}_{ji} & \tilde{z}_{jj} \end{bmatrix}\right) \end{aligned}$$

$$\begin{aligned} \operatorname{sgn}\left(\det \begin{bmatrix} z_{ii} & z_{ij} & z_{ik} \\ z_{ji} & z_{jj} & z_{jk} \\ z_{ki} & z_{kj} & z_{kk} \end{bmatrix}\right) &= \operatorname{sgn}(z_{ii} z_{jj} z_{kk} + z_{ji} z_{kj} z_{ik} + z_{ij} z_{jk} z_{ki} - z_{ik} z_{jj} z_{ki} \\ &\quad - z_{jk} z_{kj} z_{ii} - z_{ij} z_{ji} z_{kk}) \end{aligned}$$

$$= \operatorname{sgn}\left(\frac{c_v}{p_i^2}\tilde{z}_{ii}\frac{c_v}{p_j^2}\tilde{z}_{jj}\frac{c_v}{p_k^2}\tilde{z}_{kk} + \frac{c_v}{p_j p_i}\tilde{z}_{ji}\frac{c_v}{p_k p_j}\tilde{z}_{kj}\frac{c_v}{p_i p_k}\tilde{z}_{ik} + \frac{c_v}{p_i p_j}\tilde{z}_{ij}\frac{c_v}{p_j p_k}\tilde{z}_{jk}\frac{c_v}{p_k p_i}\tilde{z}_{ki} - \frac{c_v}{p_i p_k}\tilde{z}_{ik}\frac{c_v}{p_j^2}\tilde{z}_{jj}\frac{c_v}{p_k p_i}\tilde{z}_{ki} - \frac{c_v}{p_j p_k}\tilde{z}_{jk}\frac{c_v}{p_k p_j}\tilde{z}_{kj}\frac{c_v}{p_i^2}\tilde{z}_{ii} - \frac{c_v}{p_i p_j}\tilde{z}_{ij}\frac{c_v}{p_j p_i}\tilde{z}_{ji}\frac{c_v}{p_k^2}\tilde{z}_{kk}\right)$$

$$= \operatorname{sgn}\left(\left[\frac{c_v^3}{p_i^2 p_j^2 p_k^2}\right][\tilde{z}_{ii}\,\tilde{z}_{jj}\,\tilde{z}_{kk} + \tilde{z}_{ji}\,\tilde{z}_{kj}\,\tilde{z}_{ik} + \tilde{z}_{ij}\,\tilde{z}_{jk}\,\tilde{z}_{ki} - \tilde{z}_{ik}\,\tilde{z}_{jj}\,\tilde{z}_{ki} - \tilde{z}_{jk}\,\tilde{z}_{jk}\,\tilde{z}_{ii} - \tilde{z}_{ij}\,\tilde{z}_{ji}\,\tilde{z}_{kk}]\right)$$

$$= \operatorname{sgn}\left(\det\begin{bmatrix}\tilde{z}_{ii} & \tilde{z}_{ij} & \tilde{z}_{ik}\\ \tilde{z}_{ji} & \tilde{z}_{jj} & \tilde{z}_{jk}\\ \tilde{z}_{ki} & \tilde{z}_{kj} & \tilde{z}_{kk}\end{bmatrix}\right).$$

Thus, the concavity requirements (3A2.11)–(3A2.13) are equivalent to the following conditions:

$$\tilde{z}_{ii} < 0; \tag{3A2.14}$$

$$\det\begin{bmatrix}\tilde{z}_{ii} & \tilde{z}_{ij}\\ \tilde{z}_{ji} & \tilde{z}_{jj}\end{bmatrix} > 0; \tag{3A2.15}$$

$$\det\begin{bmatrix}\tilde{z}_{ii} & \tilde{z}_{ij} & \tilde{z}_{ik}\\ \tilde{z}_{ji} & \tilde{z}_{jj} & \tilde{z}_{jk}\\ \tilde{z}_{ki} & \tilde{z}_{kj} & \tilde{z}_{kk}\end{bmatrix} < 0. \tag{3A2.16}$$

It should be noted here that c_v is not identifiable from data. Although the expression of z_{ii} and z_{ij} in (3A2.7) and (3A2.9) contain this unobservable c_v, the expression of z_{ii} and z_{ij} in (3A2.8) and (3A2.10) consists of all observable (estimatable) parameters. Thus, (3A2.14)–(3A2.16) can be used to examine whether the observed 'cost function' actually satisfies the concavity requirements.

Elasticity of Substitutions

Allen–Uzawa's elasticity of substitution (AES) is defined as:

$$AES_{ij} = \frac{C_V \dfrac{\partial^2 C_V}{\partial p_i \partial p_j}}{\dfrac{\partial C_V}{\partial p_j}\dfrac{\partial C_V}{\partial p_i}}.$$

Since we have from (3A2.1), (3A2.4), (3A2.5) and (3A2.6),

$$\frac{C_V \dfrac{\partial^2 C_V}{\partial p_i \partial p_j}}{\dfrac{\partial C_V}{\partial p_j}\dfrac{\partial C_V}{\partial p_i}} = \frac{c_v \dfrac{\partial^2 c_v}{\partial p_i \partial p_j}}{\dfrac{\partial c_v}{\partial p_j}\dfrac{\partial c_v}{\partial p_i}} = \frac{\dfrac{p_i p_j}{c_v}\dfrac{\partial^2 c_v}{\partial p_i \partial p_j}}{\dfrac{p_j}{c_v}\dfrac{\partial c_v}{\partial p_j}\dfrac{p_i}{c_v}\dfrac{\partial c_v}{\partial p_i}} = \frac{\dfrac{p_i p_j}{c_v}\dfrac{\partial^2 c_v}{\partial p_i \partial p_j}}{\dfrac{\partial \ln c_v}{\partial \ln p_j}\dfrac{\partial \ln c_v}{\partial \ln p_i}} = \frac{\dfrac{p_i p_j}{c_v} z_{ij}}{\eta_i \eta_j}$$

we obtain the following neat expression of AES in our model by substituting (3A2.7) and (3A2.9) into the above expression:

$$AES_{ij} = \frac{\dfrac{p_i p_j}{c_v} z_{ij}}{\eta_i \eta_j} = \frac{1}{\eta_i \eta_j}\frac{p_i p_j}{c_v}\left(\frac{c_v}{p_i p_j}\tilde{z}_{ij}\right) = \frac{1}{\eta_i \eta_j}(\tilde{z}_{ij}) = \frac{1}{\eta_i \eta_j}(\gamma_{ij} + \eta_i \eta_j) = \frac{\gamma_{ij}}{\eta_i \eta_j} + 1.$$

Morishima's elasticity of substitution (MES) is defined as:

$$MES_{ij} = \frac{p_i \dfrac{\partial^2 C_V}{\partial p_i \partial p_j}}{\dfrac{\partial C_V}{\partial p_j}} - \frac{p_i \dfrac{\partial^2 C_V}{\partial^2 p_i}}{\dfrac{\partial C_V}{\partial p_i}}.$$

Note that we have from (3A2.1), (3A2.4), (3A2.5) and (3A2.6),

$$\frac{p_i \dfrac{\partial^2 C_V}{\partial p_i \partial p_j}}{\dfrac{\partial C_V}{\partial p_j}} - \frac{p_i \dfrac{\partial^2 C_V}{\partial^2 p_i}}{\dfrac{\partial C_V}{\partial p_i}} = \frac{p_i \dfrac{\partial^2 c_v}{\partial p_i \partial p_j}}{\dfrac{\partial c_v}{\partial p_j}} - \frac{p_i \dfrac{\partial^2 c_v}{\partial^2 p_i}}{\dfrac{\partial c_v}{\partial p_i}};$$

$$\frac{p_i \dfrac{\partial^2 c_v}{\partial p_i \partial p_j}}{\dfrac{\partial c_v}{\partial p_j}} = \frac{p_i \dfrac{c_v}{p_i p_j}\left[\gamma_{ij} + \left(\dfrac{\partial \ln c_v}{\partial \ln p_j}\right)\dfrac{\partial \ln c_v}{\partial \ln p_i}\right]}{\left(\dfrac{c_v}{p_j}\right)\left(\dfrac{\partial \ln c_v}{\partial \ln p_j}\right)} = \frac{\gamma_{ij} + \left(\dfrac{\partial \ln c_v}{\partial \ln p_j}\right)\dfrac{\partial \ln c_v}{\partial \ln p_i}}{\dfrac{\partial \ln c_v}{\partial \ln p_j}}$$

$$= \frac{\gamma_{ij} + \eta_j \eta_i}{\eta_j} = \frac{\gamma_{ij}}{\eta_j} + \eta_i;$$

and

$$\frac{p_i \dfrac{\partial^2 c_v}{\partial^2 p_i}}{\dfrac{\partial c_v}{\partial p_i}}$$

$$= \frac{p_i \dfrac{c_v}{p_i^2}\left[\gamma_{ii} + \left(\dfrac{\partial \ln c_v}{\partial \ln p_i} - 1\right)\dfrac{\partial \ln c_v}{\partial \ln p_i}\right]}{\left(\dfrac{c_v}{p_i}\right)\left(\dfrac{\partial \ln c_v}{\partial \ln p_i}\right)} = \frac{[\gamma_{ii} + (\eta_i - 1)\eta_i]}{\eta_i} = \frac{\gamma_{ii}}{\eta_i} + (\eta_i - 1).$$

Thus, we obtain:

$$MES_{ij} = \frac{\gamma_{ij}}{\eta_j} + \eta_i - \left[\frac{\gamma_{ii}}{\eta_i} + (\eta_i - 1)\right] = \frac{\gamma_{ij}}{\eta_j} - \frac{\gamma_{ii}}{\eta_i} + 1.$$

APPENDIX 3A3 THE HEURISTIC PROCEDURE FOR ESTIMATING THE TRANSLOG VARIABLE COST FUNCTION

We explain in detail the heuristic procedure used to determine what inputs are variable and to estimate the translog variable cost function parameters allowing technological change (see Section 4).

Step 1 Choice of variable inputs

Substep 1.1 (five factor inputs)

1. We take all four types of labor inputs, use machines and tools as variable inputs, and estimate equation (3.8) without period dummies. Some of the estimated γ_{ii} values are positive and statistically significant, implying that the concavity requirement is not likely to be satisfied.
2. Then, we take all four types of labor inputs, use transportation machines as variable inputs, and estimate equation (8) without period dummies. Some of the estimated γ_{ii} values are positive and statistically significant, implying that the concavity requirement is not likely to be satisfied. So, we proceed to the next substep.

Substep 1.2 (four factor inputs)

1. We drop all capital stocks except for ICT stocks. We then take all four types of labor inputs as variable inputs and estimate equation (3.8) without period dummies. Some of the estimated γ_{ii} values are positive and statistically significant, implying that the concavity requirement is not likely to be satisfied.

2. We keep machines and tools and drop one of the three labor inputs (young with high levels of education, old with low levels of education,

and old with high levels of education). We estimate equation (3.8) without period dummies. Some of the estimated γ_{ii} values are positive and statistically significant, implying that the concavity requirement is not likely to be satisfied.

3. Then, we keep transportation machines and drop one of the three labor inputs. We estimate equation (3.8) without period dummies. Some of the estimated γ_{ii} values are positive and statistically significant, implying that the concavity requirement is not likely to be satisfied. So we proceed to the next substep.

Substep 1.3 (three factor inputs)

We drop all capital stocks except for ICT stocks and keep young workers with low levels of education. We then drop one of the remaining three labor inputs. We estimate equation (3.8) without period dummies. We next examine whether all estimated γ_{ii} values are negative with some statistical significance. In this step, six out of 16 industries show all negative γ_{ii} values with marginal statistical significance: food, textile, fabricated metal, electrical machinery, instruments, and services.[39] These six industries are likely to have sharper results if we consider period dummies explicitly based on technological change. Thus, we move to step 2 for these industries. For the remaining ten industries, we proceed to the next substep.

Substep 1.4 (two factor inputs)

We drop all capital stocks except for ICT stocks and keep young workers with low levels of education, drop two of the remaining three labor inputs, and estimate equation (3.8) without period dummies. We examine whether estimated γ_{ii} values are negative with some statistical significance. Out of the remaining ten industries, nine have estimated γ_{ii} values that are all negative with marginal statistical significance. We move to step 2 for these industries for the same reason given in substep 1.3. We proceed to the next substep for the general machinery category.

Substep 1.5 (period difference)

The failure of substeps 1.1–1.4 for the general machinery industry suggests that there may be a break in the number of quasi-fixed factors for that industry between the 1980s and the 1990s. We divide the total sample period into two periods and reapply substeps 1.1–1.4 for each subperiod. The results suggest that the general machinery industry has two variable factors in the 1980s and three variable factors in the 1990s. We then proceed to step 2.

Step 2 Estimation of share equations with period dummies

Substep 2.1 (identifying the possible number of technological changes)
For each industry, we identify the possible number of technological changes. As explained in the text, there may be a technological change between the 1980s and the 1990s because the usage of information technology is different in the two periods. (In the case of the general machinery industry, we found in step 1 that the number of quasi-fixed factors is different in the 1980s compared with that in 1990s, already implying a break in production technology.) In addition, there may be an additional technological change for specific industries. To identify technological changes for each industry, we look at Table 3.10, which reports ICT stocks' contribution to value-added growth for the entire sample period and for each half-decade. For some manufacturing industries, the ICT capital stock contribution has a break in the late 1990s, which may indicate another technological change. In the case of the transportation and communication sector, we find a sharp increase in the ICT contribution in the 1980s. This suggests that there might be a technological change in the mid-1980s in this industry. Taking these observations into account, we consider additional intercept and slope dummies of the mid-1990s for manufacturing industries having a break in ICT contribution within the 1990s. In the case of the transportation and communication industry, we consider the mid-1980 dummies instead of the mid-1990 dummies.

Substep 2.2 (searching for the timing of technological change)
For the technological change between the 1980s and 1990s, we first set 1990 as the year of change. For the industry-specific technological change (suggested in the previous substep), we set 1995 for the change in the 1990s and 1985 for the change in the 1980s. Upon deciding the number of period dummies (that is, technological changes), we estimate equation (8) with these period dummies, drop insignificant intercept and/or slope dummies, re-estimate the equations, and examine whether estimated coefficients are consistent with the concavity requirement. We then move the point of change around the initial point to see whether this gives us a sharper estimation (in terms of the statistical significance of the γ_{ii} values), with the concavity requirement remaining satisfied. In the end, some period dummies are not statistically significant. The results are reported in Tables 3.3a and 3.4a.

Step 3 Manufacturing industries

For manufacturing industries, we have one more dimension with respect to labor input types: production workers and nonproduction workers. This means we have eight types of labor inputs. Basically, we repeat steps 1 and 2 for these finer labor input data of manufacturing. There is a problem of apparent multicollinearity, however, and consequently some form of aggregation is necessary. We try all sensible aggregation possibilities and find that aggregating young and old production workers and using six types of labor inputs (production workers with low education levels, production workers with high education levels, nonproduction young workers with low education levels, nonproduction young workers with high education levels, nonproduction old workers with low education levels, and nonproduction old workers with low education levels) yielded satisfactory results. The results are reported in Tables 3.3b and 3.4b.

NOTES

* This chapter is an augmented version of: Kiyohiko G. Nishimura and Masato Shirai (2003), 'Can information and communication technology solve Japan's productivity-slowdown problem?', *Asian Economic Papers*, 2 (1), 85–139. We are grateful to past discussants and participants of the ESRI conferences, and to Kaliappa Kalirajan and Jong-Wha Lee of the Asian Economic Panel for helpful comments. We are also grateful to Tsutomu Miyagawa of Gakushuin University and Japan Center for Economic Research, who allowed us to use their disaggregated capital stock series. The work reported here is partially supported by a grant from the Economic and Social Research Institute, the Cabinet Office, the Government of Japan. The views expressed here are the personal views of the authors and in no way represent the views of the institutions the authors belong to, or those of the Economic and Social Reearch Institute.

1. A notable exception is Hiromatsu et al. (2000).
2. The efforts of recent governments to promote ICT have clearly been based on this belief (for example, the e-Japan program of the Koizumi government).
3. There is now a sizeable literature on this effect in Japanese industries (most studies are written in Japanese). See Nishimura and Morita (2002) and the references therein.
4. We cannot state a priori that all capital stocks are quasi-fixed and that all labor inputs are perfectly variable. In this respect, our approach is different from the quasi-fixed-capital literature (for example, Morrison 1992; Flaig and Steiner 1993), which does assume a priori that all capital goods are quasi-fixed and all labor inputs are variable. Our versatile setting, however, obliges us to assume homotheticity and constant returns to scale, whereas the approaches of Morrison, and Flaig and Steiner do not.
5. Examining a large panel of firms, these authors find non-negligible deviation from perfect competition in almost all industries in Japan.
6. Our investigation focused on within-industry ICT externality, rather than cross-industry externality. Stiroh (2001) obtains the same result for his compilation of detailed industry-wise data in the United States. There is a sizeable literature claiming to show a positive correlation between ICT investment and total factor productivity growth, based on aggregate data (see, for example, Van Ark 2001; Haacker and Morsink 2002; and Economic Planning Agency 2000). Their findings of the literature may show that there is cross-industry ICT externality. These studies rely on rather crude aggregate data, however, and make some heroic assumptions, such as setting the elasticity of output with respect to capital at 0.35 for all countries and for all sample periods under investigation (Haacker and Morsink 2002). Thus, although these studies are very illuminating first attempts on this subject, further careful disaggregate analysis is required to settle the question of whether cross-industry ICT externality exists.

7. See Appendix 3A1.
8. It is straightforward to extend our analysis to the case in which some quasi-fixed factor inputs must be determined well in advance before production, though the notation becomes cumbersome. For example, consider the case of two quasi-fixed factor inputs. The following analysis does not change if one factor must be determined, say, two periods before production, whereas the other factor is determined one period before production, so long as the production capacity function is multiplicatively separable, such that $S = S^1(z_1)S^2(z_2)$, where S^1 and S^2 are homogeneous of degree k' and k'' and $k' + k''' = 1 + k$. We then have a three-period sequential expected profit maximization to determine z_1 and z_2, instead of a two-period expected profit maximization, as described above.
9. We do not use the term total factor productivity (TFP) here. Precisely speaking, TFP is defined as the ratio of the Divisia index of outputs to that of inputs. Although TFP growth is equal to the rate of technological progress if all factors are variable under perfect competition and constant returns to scale, this relationship does not hold in more general cases. We do not assume that all factors are variable, nor do we assume that competition is perfect.
10. The meaning of this 'relevant period' will be explained in Section 5.
11. We are now starting a project that examines possible biases that the definitional deviation has on GDP and TFP growth analysis.
12. In the case of IT software, only the 1995 fixed capital formation matrices of the *Base-Year Input Output Tables* report industry-by-industry data. We extrapolate the series before 1995 by using the METI's *Information Technology Survey* (1980–1998).
13. The Bureau of Research of the Economic Planning Agency followed a similar procedure in its *Policy Effectiveness Analysis Report No. 4*, October 2000.
14. Schreyer's work analyses the contribution of ICT to output growth in the G7 countries and uses the same definition that we do for ICT hardware. An alternative deflator is the wholesale price index of ICT hardware products that is published by the Bank of Japan. This price index has serious problems, however, which are described in the following paragraphs. Schreyer's data ended at 1996, so we extrapolated his data into 1997 and 1998 by means of a linear trend.
15. The only true ICT software investment deflator is reported in the corporate service price index (also compiled by the Bank of Japan), which is a price index of software development. This index is available only for three years (1995–98). Thus, we are obliged to use the imputation method described in the text.
16. Tax data were not available for 1998, so we substituted 1997 data for the 1998 figures.
17. There are three kinds of employed workers: employees, self-employed workers and family workers. The *Basic Survey on Wage Structure* contains information only for employees. Thus, we supplement these data with figures from the *Annual Report on the Labor Force Survey*, produced by the Statistics Bureau of the Management and Coordination Agency, which contains information about self-employed workers and family workers. This publication does not break down self-employed and family workers into the various subcategories we consider or provide information that would enable us to do so, so we postulate that the breakdown is the same as that for employees. Kuroda et al. (1997) also assumed this breakdown.
18. For employees, the *Basic Survey on Wage Structure* has work-hour information. For self-employed and family workers, we use the *Annual Report on the Labor Force Survey*.
19. The wage income data in the *Basic Survey on Wage Structure* do not correspond exactly to the compensation data based on the System of National Accounts (SNA). Thus, we first estimate wage payments for each worker type using the *Basic Survey on Wage Structure* data, then divide the SNA total compensation of employees into compensation for each worker type, relying on this obtained distribution of wage income.
20. For self-employed workers and family workers, we adopt the method of Kuroda et al. (1997). See their unpublished appendix for details.
21. In fact, we applied the same procedure to these excluded industries alongside the other industries. We found that these industries' estimated cost functions did not satisfy the

concavity requirement or even the monotonicity requirement, no matter what set of inputs was chosen as variable inputs.

22. We applied the same procedure to the 1990s finance industry data, but the estimated cost functions did not satisfy the concavity requirement nor even the monotonicity requirement, no matter which inputs were chosen as the variable inputs.
23. Stiroh (2001) classified an industry as ICT-intensive if the ratio of ICT capital stocks to total capital stocks in that industry was above the average ratio for all industries in 1985–89, just before the acceleration of ICT investment in the 1990s. Using the same criterion, we identify five ICT-intensive industries.
24. To our knowledge, this is the first attempt of this kind. Suruga and Hashimoto (1996) survey several attempts to discern substitutability or complementarity between various labor inputs and capital stocks. None of the studies Suruga and Hashimoto surveyed, however, examined substitutability and complementarity between ICT stocks and labor inputs.
25. See Appendix 3A2.
26. In general, the concavity requirement on cost functions is not neatly represented by restrictions on the parameters β_i and γ_{ij}. Thus, the share function is customarily estimated by imposing only homogeneity of degree one, and then one determines whether the estimated parameters imply the concavity of the cost function locally around the sample mean of input prices.
27. Here we again take a heuristic approach: we evaluate terms at the average input price vector of the relevant period and examine whether sign conditions of concavity are satisfied or not. If sign conditions are satisfied at the average price vector, we consider the estimated cost function to satisfy the concavity requirement. Though not reported here, we have previously tested whether the concavity requirement is met each year. The result suggests that concavity is satisfied in most industries and in most subperiods, with the possible exception of the fabricated metal and textile industries in the 1980–84 subperiod. (This statement is based on the lower bound of probability that the concavity requirement is met. Here the concavity requirement is represented by a set of joint hypotheses on principal minors of the Hessian matrix.)
28. By 'small deviation' we mean that some of the sign conditions may be violated, but the deviation is rather small in absolute terms.
29. 'It is standard in the literature to define the level of labor skill on the basis of the level of workers' education' (Krusell et al. 2000, p. 1033).
30. We exclude finance and insurance industry because data for this industry after 1995 are problematic.
31. It should be noted here that we include software output in GDP, whereas the published GDP figure, which is based on the 1968 SNA convention, does not. In the published figure, software is considered to be an intermediate input rather than a final-goods output.
32. Many microeconomic studies find a large economic impact from ICT use in firms. See the surveys of Brynjolfsson and Yang (1996) and Brynjolfsson and Hitt (2000). In addition, recent studies on this subject based on aggregate data find evidence of ICT externality of this kind (see note 6 and references therein).
33. Stiroh (1999a) reviews the New Economy literature.
34. See Gordon (1990) and De Long and Summers (1992).
35. We tried other specifications, including one allowing a lag structure using annual data, but the result was not promising. In particular, we tried a specification that Jong-Wha Lee suggested, allowing a technology spillover effect of ICT through high-skilled workers (a cross-term of ICT and high-education labor). Unfortunately, a severe multicollinearity problem was found, ruining the regression results.
36. Instruments we use are (1) constant; (2) 90sDummy; (3) the ratio of the old (40 years and over) in the total population; (4) the ratio of college and junior college graduates in the total 20- to 24-year-old population of 1951–55, 1956–60, 1961–65 and 1966–70; (5) population growth; (6) one-period-lagged value-added growth; (7) one-period-lagged capital/labor ratio; (8) one-period-lagged ITK estimated using Miyagawa's ICT capital;

and (9) one- period-lagged EQ. Hansen's overidentifying restrictions test (Hansen 1982) shows that our choice is reasonable. We also tried other macroeconomic variables, but the results were not satisfactory because of apparent multicollinearity.

37. See Nishimura and Tamai (2001) for a model of long-run rigidity of management styles.
38. See Nishimura and Morita (2002) for the importance of product architecture in understanding industrial structure. Japan is still highly competitive in the automobile industry, since the latter is an integral-architecture industry.
39. Moreover, the combination of chosen labor inputs is unique to each industry in this category. That is, there is only one combination of labor inputs for each industry that has all negative γ_{ii} values with some statistical significance.

REFERENCES

Ariga, Kenn, Yasushi Ohkusa and Kiyohiko G. Nishimura (1999), 'Determinants of individual firm mark-up in Japan: market concentration, market share and FTC's Regulations', *Journal of Japanese and Industrial Economies*, **13** (4), 424–50.

Blackorby, Charles and R. Robert Russell (1989), 'Will the real elasticity of substitution please stand up? (A Comparison of the Allen/Uzawa and Morishima Elasticities)', *American Economic Review*, **79** (4), 882–8.

Brynjolfsson, Erik and Lorin Hitt (2000), 'Beyond computation: information technology, organizational transformation and business practices', *Journal of Economic Perspectives*, **14** (4), 23–48.

Brynjolfsson, Erik and Shinkyu Yang (1996), 'Information technology and productivity: a review of the literature', *Advances in Computers*, **43**, February, 179–214.

Bureau of Economic Analysis (BEA), US Department of Commerce (1997), *Survey of Current Business*, Washington, DC: US Government Printing Office.

Bureau of Research (2000), *Policy Effectiveness Analysis Report No. 4*, October, Tokyo: Economic Planning Agency.

De Long, J. Bradford and Lawrence Summers (1992), 'Equipment investment and economic growth: how strong is the nexus?', *Brookings Papers on Economic Activity*, **2**, 157–99.

Economic Planning Agency (2000), *Keizai-Hakusho* (White Paper on Economic Activities). Tokyo: Economic Planning Agency.

Flaig, Gebhard and Viktor Steiner (1993), 'Searching for the 'productivity slowdown': some surprising findings from West German manufacturing', *Review of Economics and Statistics*, **75** (1), 57–65.

Fraumeni, Barbara M (1997), 'The measurement of depreciation in the U.S. National Income and Product Accounts', *Survey of Current Business*, **77** (7), 7–23.

Gordon, Robert J. (1990), *The Measurement of Durable Goods Prices*, National Bureau of Economic Research Monograph, Chicago: University of Chicago Press.

Haacker, Markus and James Morsink (2002), 'You say you want a revolution: information technology and growth', IMF Working Paper WP/02/70, Washington, DC: International Monetary Fund.

Hansen, Lars (1982), 'Large sample properties of generalized method of moments estimators', *Econometrica*, **50** (4), 1029–54.

Hiromatsu, Takeshi, Manabu Kurita, Naoki Tsbune, Minoru Kobayashi and Gohsei Ohira (2000), 'Quantitative analysis on labor-saving effect of informatization' (in Japanese), *Keiei Jouhou Gakkai Shi*, **8** (4), 49–65.

Institute of Local Finance (1980–1997), *Annual Statistical Report of Local Governments* (*Chihou Zaisei Toueki Nennpou*) (in Japanese), Tokyo: Printing Bureau, Ministry of Finance, Japan.

Jorgenson, Dale W. and Kevin J. Stiroh (2000), 'Raising the speed limit: U.S. economic growth in the information age', *Brookings Papers on Economic Activity*, **1**, 125–236.

Krusell, Per, Lee E. Ohanian, José-Víctor Ríos-Rull and Giovanni L. Violante (2000), 'Capital–skill complementarity and inequality: a macroeconomic analysis', *Econometrica*, **68** (5), 1029–53.

Kuga, Kiyoshi (1979), 'On the symmetry of Robinson elasticities of substitution: the general case', *Review of Economic Studies*, **46** July, 527–31.

Kuroda, Masahiro, Kazushige Simpo, Koji Nomura and Nobuyuki Kobayashi (1997), *KEO Data Base – The Measurement of Output, Capital and Labor* (in Japanese), Monograph Series no. 8, Tokyo: Keio Economic Observatory.

Ministry of Economy, Trade, and Industry (METI) (1980–1998), *Information Technology Survey* (*Wagakuni Jyouhousyori no Genjyou*) (in Japanese), Tokyo: Printing Bureau, Ministry of Finance, Japan.

Ministry of Finance, Policy Research Institute (1980–1998), *Financial Statements Statistics of Corporations* (*Houjinn Kigyou Toukei*) (in Japanese), Tokyo: Printing Bureau, Ministry of Finance, Japan.

Ministry of Health, Labor and Welfare (1980–1998), *Basic Survey on Wage Structure* (*Chinnginn Kouzou Kihonn Toukeichyousa*) (in Japanese), Tokyo: Rodo Horei Kyokai, Japan.

Ministry of International Trade and Industry (MITI) (1980, 1985, 1990, 1995), *Base-year Input Output Tables*, Tokyo: Statistical Standards Department, Statistics Bureau, Management and Coordination Agency, Japan.

Miyagawa, Tsutomu, Yukiko Ito and Nobuyuki Harada (2001), 'Industry-wise IT investment and its spillover effects', in Japan Center for Economic Research (ed.), *Firm Behavior in the Information Age* (in Japanese), Tokyo: Japan Center for Economic Research, Division of Economic Research, pp. 41–71.

Miyagawa, Tsutomu and Suyuri Shiraishi (2000), 'Why did economic growth decline in Japan? An analysis using new capital stocks data categorized by asset' (in Japanese), JCER Discussion Paper no. 62, Tokyo: Japan Center for Economic Research.

Morrison, Catherine J. (1992), 'Unraveling the productivity slowdown in the U.S., Canada, and Japan: the effects of subequilibrium, scale economies and markups', *Review of Economics and Statistics*, **74** (3), 381–93.

Murota, Takeshi (1977), 'On the symmetry of Robinson elasticity of substitution: a three-factor case', *Review of Economic Studies*, **44** (1), 173–6.

National Tax Agency (1980–1998), *Survey on Corporate Activities* (*Houjinkigyou no Jittaityousa*) (in Japanese), Tokyo: Printing Bureau, Ministry of Finance, Japan.

Nishimura, Kiyohiko G. and Masataka Morita (2002), 'Information technology and automobile distribution: a comparative study of Japan and the United States', *International Journal of Automobile Technology Management*, **2** (2), 206–37.

Nishimura, Kiyohiko G., Yasushi Ohkusa and Kenn Ariga (1999), 'Estimating the mark-up over marginal cost: a panel analysis of Japanese firms 1971–1994', *International Journal of Industrial Organization*, **17** (8), 1077–111.

Nishimura, Kiyohiko G. and Masato Shirai (2000), 'Fixed cost, imperfect competition and bias in technology measurement: Japan and the United States', Discussion Paper no. 273, Washington, DC: OECD, http://www.oecd.org/pdf/M00002000/M00002108.pdf.

Nishimura, Kiyohiko G. and Yoshihiro Tamai (2001), 'Long-run rigidity in labor markets', in Economic Theory, Dynamics and Markets, Takashi Negishi, Rama Van Ramachadran and Kazuo Mino (eds), Boston, MA: Kluwer Academic, pp. 93–103.

Oliner, D. Steven and Daniel E. Sichel (2000), 'The resurgence of growth in the late 1990s: is information technology the story?', *Journal of Economic Perspectives*, **14** (4), 3–22.

Schreyer, Paul (2000), 'The contribution of information and communication technology to output growth: a study of the G7 countries', STI Working Paper 2000/2, Washington, DC: OECD.

Statistics Bureau, Management and Coordination Agency, Government of Japan (1980–1998), *Annual Report on the Labor Force Survey* (*Roudouryoku Chousa Nennpou*) (in Japanese), Tokyo: Japan Statistical Association.

Stiroh, Kevin J. (1999a), 'Is There a New Economy?', *Challenge*, **42** (4), 82–101.

Stiroh, Kevin J. (1999b), 'Measuring input substitution in thrifts: Morishima, Allen–Uzawa, and cross-price elasticities', *Journal of Economics and Business*, **51** (2), 145–57.

Stiroh, Kevin J. (2001), 'Are ICT spillovers driving the new economy?', *Review of Income and Wealth*, 48 (1), 33–58.

Suruga, Terukazu and Keiji Hashimoto (1996), 'On the substitution between labor differentiated by educational attainment and capital: the case of Japanese manufacturing industries' (in Japanese), *Journal of the Japan Statistical Society*, **26** (3), 255–67.

Van Ark, Bart (2001), 'The renewal of the old economy: an international comparative perspective', STI Working Paper 2001/5, Washington, DC: OECD.

4. Demand saturation – creation and economic growth

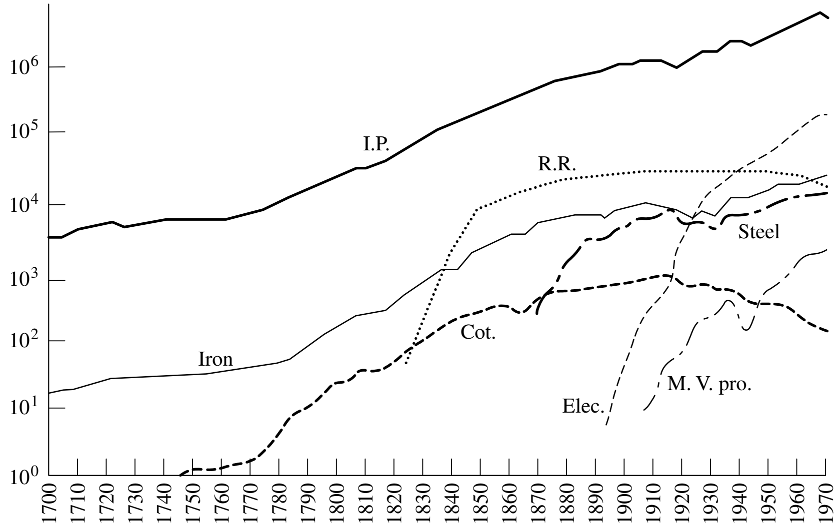
Masanao Aoki and Hiroshi Yoshikawa*

1. INTRODUCTION

In the standard literature, the fundamental factor restraining economic growth is diminishing returns to capital in production or research and development (R&D) technology. In this chapter, we present a model suggesting that 'saturation of demand' is another important factor restraining growth.

In the less mathematical literature and casual discussions, the idea of 'demand saturation' has been very popular. Every business person would acknowledge saturation of demand for an individual product. In fact, plot a time series of production of any representative product such as steel and automobiles, or production in any industry, against year, and, with few exceptions, one obtains an S-shaped curve. Figure 4.1, from Rostow (1978), demonstrates this 'stylized fact'. The experiences of diffusion of such consumer durables as refrigerators, television sets, cars and personal computers tell us that deceleration of growth comes mainly from saturation of demand rather than diminishing returns in technology. Growth of production of a commodity or in an individual industry is bound to slow down because demand grows fast at the early stage but eventually, necessarily slows down. Thus the demand for some products grows much more rapidly than the gross domestic product (GDP), while the demand for others grows much more slowly. Products/industries face different income elasticities of demand. The celebrated Engel's law, based on saturation of demand for food, is merely an example.

Unfortunately, the existing literature on growth abstracts largely from this important fact that products/industries obey the law of demand saturation and that each product/industry experiences a typical S-shaped life cycle. This, of course, is not to say that the appearance of new products and the disappearance of old ones have not been modeled. The so-called 'creative destruction' and the 'quality ladder' literature, such as Grossman and Helpman (1991), Aghion and Howitt (1992), Caballero and Jaffe (1993)



Notes: Industrial production (I.P.); railroad mileage (R.R.); raw cotton consumption (Cot.); production of iron (Iron); steel (Steel); electricity (Elec.); and motor vehicles (M.V. Pro.).

Source: Rostow (1978).

Figure 4.1 Aggregate and sectorial growth patterns illustrated: British industrial production and seven major sectors, 1700–1972

and Horowitz and Lai (1996), has analysed such phenomena in growth models. However, in this line of research, the old products disappear only through the introduction of new products. Unless new products appear, demand for the existing products remains the same. Therefore, it is possible for the economy to keep growing if it succeeds in raising productivity in the production of the existing commodities.[1]

In contrast, with saturation of demand as we assume it in the present chapter, to raise productivity in production of the 'mature' products does not help in sustaining economic growth. To put it another way, in the existing R&D-based growth models, the economy can keep growing, if, for example, the automobile industry keeps raising the quality of cars, whereas in our model, it cannot because demand for cars saturates in spite of quality improvement.

Likewise, the product life cycle in the existing literature such as Grossman and Helpman (1991) is based on a production technology or production geography life cycle, while in our model it is based on a demand life cycle. In contrast to the creative destruction that occurs in the existing literature, growth of demand for the existing commodities in our analysis of saturation necessarily slows downs whether or not new commodities appear. It would be absurd to argue that the growth of demand for food decelerated, as Engel found, because manufactured products appeared. The demand for cars did not approach its ceiling because personal computers were invented. Rather, the law of demand saturation works for an individual commodity.

Within the same industry, new and old products are often close substitutes like black-and-white and color TVs or personal computers of different vintage, and old products gradually disappear as new ones appear. Thus, the creative destruction story nicely fits the growth of an industry. The R&D race among competing firms as it is modeled in the standard endogenous growth literature certainly plays an important role. Technical progress taken up in the existing literature basically concerns close substitutes, as those models explicitly state. However, as we argued above, the same story does not necessarily hold true for different industries. Arguably, demand saturation is more relevant for the growth of the economy as a whole. In this chapter, we explore the growth model based on demand saturation and innovations which create new demand. To repeat, in the standard 'creative destruction', 'quality ladder' or 'product variety' models, the economy can sustain growth if productivity in the population of existing products keeps rising, while in our demand saturation model, it cannot.

We take the logistic growth of an individual product/industry as a stylized fact, and present a formal model of growth built on this stylized fact. An obvious implication of the logistic growth of an individual product/industry is that the economy enjoys high growth if it successfully keeps introducing new products or industries which temporarily enjoy *high growth of demand.* In this chapter, innovation or 'technical progress' bears new commodities or sectors which enjoy high growth of demand and, by so doing, sustains the economic growth of the economy as a whole.

The demand-creating innovation in our model is different from the standard total factor productivity (TFP), or an 'upward shift' of the production function. In the standard quality ladder models and the creative destruction literature such as Grossman and Helpman (1991), Caballero and Jaffe (1993), Horowitz and Lai (1996), and Young (1998), innovation or technical progress basically raises TFP by way of replacing old commodities with new ones simply because new commodities are assumed to have higher value than old ones. Again, whereas this seems to hold true for the commodities which are basically the same but of a different vintage, the same story does not make much sense for wholly different products such as cars and personal computers. Personal computers do not necessarily command higher value added than cars. In short, the standard literature models the dynamics of close substitutes while our model stresses the importance of demand saturation and creation of wholly different products or industries for which demand grows transitionally.

The difference between the standard models and our model of demand saturation most clearly shows up in the transitory dynamics. In the standard R&D-based growth models, the efficiency of R&D determines the transitory dynamics, whereas in our model, the pattern of demand saturation is the determinant. As a model of demand-constrained growth, our model follows the long line of post-Keynesian literature such as Kaldor (1957), Robinson (1962), Marglin (1984) and Dutt (1986). In the post-Keynesian tradition, income distribution between capital and labor plays a central role in determining aggregate demand. In contrast, our analysis focuses on saturation of demand for an individual good/sector as a factor to generate demand constraints facing the economy. We discuss the transitory dynamics in Section 3.

Innovations in the economy facing the law of demand saturation contribute to growth in a different way from an upward shift of the production function does. That TFP does not necessarily capture the significance of technological progress is pointed out by Wright (1997, p. 1562):

> The identification of 'technological progress' with changes in total-factor-productivity, or with the 'residual' in a growth-accounting framework, is so widely practised that many economists barely give it a passing thought, regarding the two as more-or-less synonymous and interchangeable. . . . Even with extensive quality adjustments, TFP is not generally a good index of technology. If a genuine change in technological potential occurs in a firm, an industry, a sector, or a country, in any plausible model this change will affect the mobilisation of capital and labour in whatever unit is involved. In the new equilibrium, inputs as well as outputs will have changed; the ratio between these may convey little if any useful information about the initiating change in technology.

We share Wright's concern. The economy always mobilizes resources and accumulates capital whenever it finds goods or sectors for which demand grows rapidly. In fact, in our model, the elasticity of capital in the production function is equal to one (the so-called AK model). Therefore, the economy grows whenever capital accumulates. But capital accumulation is constrained by saturation of demand. Innovation creates goods/sectors for which demand grows fast, elicits capital accumulation and thereby ultimately sustains economic growth.

To substantiate this argument, in Section 2, we present a model which incorporates the basic idea that demand for an individual good or sector necessarily faces saturation and thus its growth eventually slows down. We begin with demand for an individual product rather than preferences, since the former is more directly related to the stylized fact than the latter. Section 3 studies growth of the economy as a whole. Out of the steady state, 'vigor of demand' and saturation determine growth while the ultimate factor for sustaining economic growth in the steady state is the creation of new products/industries. Under the standard Poisson assumption, successive creation of new products/industries sustains steady-state growth. However, we demonstrate that under the alternative 'Polya urn' assumption, the success probability of innovation gets smaller as time goes by, the growth rate of the economy must decelerate and go asymptotically down to zero. This is the same result as that obtained in the standard R&D-based TFP models such as Jones (1995), Jones and Williams (1998), Segerstrom (1998) and Young (1998), though the rate of innovation is a decreasing function of time rather than a function of the R&D capital stock, as is assumed in the existing literature. Section 4 provides microeconomic foundations for investment and consumption. For consumption which follows the logistic growth, we present two different microeconomic foundations: Ramsey model with the representative consumer, and a model with diffusion of goods among different households. The two models suggest different interpretations of saturation of demand. Finally, Section 5 offers some concluding remarks.

2. THE MODEL

We study an economy in which heterogeneous final goods and an intermediate good are produced. In this section, we take demand for each final product as given, and concentrate on production. We shall later consider the firm's behavior which determines investment and also the consumer's behavior which determines consumption in Section 4. Let us begin with final goods.

Final Goods

Final goods are produced with an intermediate good as the only input. Production of all the final goods requires the same intermediate good X. The production function is also common:

$$y_k = AX_k \quad (0 < A < 1). \tag{4.1}$$

We assume perfect competition. Therefore, zero profits ensue:

$$P_k(t)y_k(t) = P_X(t)X_k(t). \tag{4.2}$$

Here $P_k(t)$ is the price of the kth final product, and $P_X(t)$ the price of intermediate good.

Because of the common linear production function (4.1), the zero profit condition (4.2) is equivalent to:

$$P_k(t)A = P_X(t). \tag{4.3}$$

Thus we can adjust the units of final products in such a way as to make all the prices of final goods equal to 1. Then: $P_X = A < 1$.

Output of each final good is equal to its demand $D_k(t)$ no matter how the latter is determined:

$$y_k(t) = D_k(t). \tag{4.4}$$

In this section, we take an S-shaped life cycle of demand for each product/industry as a stylized fact. To make our analysis tractable, assume that $D_k(t)$ follows the logistic curve:

$$D_k(t) = \frac{\mu D_0}{[\delta D_0 + (\mu - \delta D_0)e^{-\mu t}]}. \tag{4.5}$$

Since the mechanism is the same for all the products or sectors, for the moment, we drop k and write $D_k(t)$ as $D(t)$. We will explore microeconomic foundations for the logistic growth of demand in Section 4. D_0 in (4.5) is the initial value of $D(t)$. Starting with D_0 smaller than μ/δ, $D(t)$ initially increases almost exponentially, but its growth eventually decelerates, and approaching its 'ceiling' μ/δ, the growth rate declines asymptotically to zero. A typical shape of the logistic growth is illustrated and compared with exponential growth in Figure 4.2.

Though exponential growth is often taken for granted by economists, there is actually ample evidence to show that no *individual* product or industry grows exponentially. Rather, demand for, or production of, a product or an industry typically grows according to the logistic curve. In fact, an eminent mathematician, Montroll (1978), goes so far as to suggest that almost all the social phenomena, except in their relatively brief abnormal times, obey the logistic growth. Figure 4.1 demonstrates this well-known fact of life in our economy.

Growth and saturation of an individual product/sector are here characterized by two parameters μ and δ, which would depend not only on preferences but also on creation of new models and close substitutes of higher quality in the production of the same commodity. For example, TV set growth would certainly have reached its ceiling much earlier if there had been only black-and-white TVs; the emergence of color TVs and small

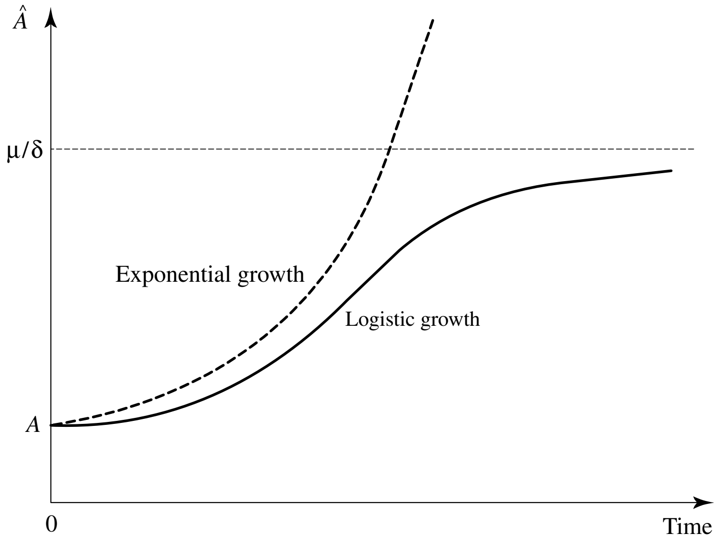


Figure 4.2 The logistic curve

models pushed up the ceiling. We maintain, however, that such technical progress cannot overcome the law of demand saturation in the end.[2] In fact, based on his careful study of the US patent data, Schmookler (1966) even argues that technical progress in one industry is itself very strongly conditioned on the prospects of demand in that industry.

Following the logistic growth of demand, production $y_k(t)$ also satisfies equation (4.5). So far, we have focused on a final good. The number of final products is not given, however. Rather, at any moment a new product or sector may arise, and the emergence of an utterly new final good or a new sector is the result of innovations. Before we explain this, we turn to production of intermediate goods taking the number of final goods N as if it were constant.

Intermediate Good

To keep our model as simple as possible, we assume that there is only one kind of intermediate good X, and that X is produced by using capital K alone:

$$X = aK. \tag{4.6}$$

Here X is the sum of intermediate goods used in production of final goods:

$$X = \sum_{k=1}^{N} X_k.$$

The capacity utilization a is determined by the firm together with capital accumulation. We shall discuss the firm behavior in Section 4. For the moment, we can imagine that a is constant, which is true in the steady state.

We note that the production function (4.6) has a unitary elasticity of capital, and therefore that as long as capital accumulates, X grows without limit. And given the common production function for final goods (4.1), whenever X grows, production of final goods can also grow. However, X is an intermediate good and, as seen previously, the growth of demand for each final good decelerates and declines eventually to zero. In this model, the factor that limits capital accumulation and thereby growth is not diminishing returns on capital but declining growth of demand. We shall consider how the profit-maximizing firm determines the capital accumulation and the capacity utilization in Section 4.

Emergence of New Final Goods or Industries

So far we have taken the number of final goods as if it were constant. In fact, new final goods and/or industries emerge as a result of innovations. We can flexibly interpret final 'goods' as 'sectors' or 'industries' if we wish.

Much effort has been made to explicitly analyse R&D activities and inventions in growth models. In fact, the achievement of the 'endogenous growth theory' is to have combined growth models with models of R&D activities.However, as we pointed out above, technical progress in the existing literature basically concerns quality improvement in the production of *close substitutes*. Therefore, it more closely applies to an industry than to the economy as a whole. We maintain that such technical progress pushes up the ceiling of demand for the existing products, but cannot overcome the law of demand saturation in the end. Innovations in this model, in contrast, bear wholly new products/sectors/industries for which demand grows fast. Such innovations would depend not only on profit-motivated R&D but also on basic research. In any case, our primary interest is not in microeconomic foundations for R&D activities but in the way in which technical progress or innovation affects the economy.[3]

Specifically we assume that an invention of a new final good or an emergence of a new sector stems stochastically from learning in the process of production of the existing products. To be specific, we assume that the probability that a new final good is invented or a new industry emerges between t and $t + \Delta t$ is $\lambda N \Delta t$ where N is the number of existing final goods

($\lambda > 0$). Since an invention or new sector is a branch off from an existing good or sector, the rate of success probability is proportional to the number of existing final goods/sector N; the more products or sectors in the economy, the more likely a new product or sector will emerge. λ is a parameter that represents the strength of innovations or more precisely the probability that a new good or industry emerges in the existing process of production. Innovations are thus accidental, but depend on the prior 'knowledge' and experiences which stem from the existing production.

Given this assumption, $Q(N,t)$ the probability that the number of final goods at time t, $N(t)$ is equal to N, satisfies the following equation:

$$\frac{dQ}{dt} = -\lambda N Q(N,t) + \lambda(N-1)Q(N-1,t). \tag{4.7}$$

Without loss of generality, we can assume that the initial number of final goods is 1:

$$Q(N,0) = \delta(N-1).$$

Appendix 4A shows that the solution of this equation under this initial condition is:

$$Q(N,t) = e^{-\lambda t}(1 - e^{-\lambda t})^{N-1}. \tag{4.8}$$

The probability that there are N goods at time t and the $(N+1)$th good emerges between t and $t + \Delta t$ is then given by:

$$\lambda N Q(N,t)\Delta t = \lambda N e^{-\lambda t}(1 - e^{-\lambda t})^{N-1}\Delta t. \tag{4.9}$$

At time t, the production of final goods which emerged at $\tau(\tau < t) y_\tau(t)$ has grown to:

$$y_\tau(t) = \frac{\mu}{\delta + (\mu - \delta)e^{-\mu(t-\tau)}} \tag{4.10}$$

since the growth of $y_\tau(t)$ obeys the logistic curve. Again, without loss of generality we can assume the initial production of newly invented good D_0 to be 1 in equation (4.5). Before we provide microfoundations for consumption and investment, we consider the growth of the economy as a whole.

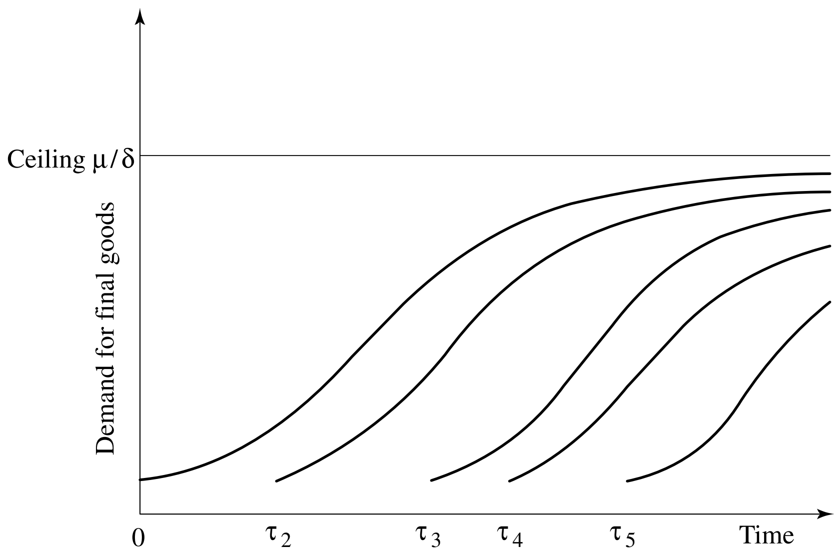
3. GROWTH OF THE MACROECONOMY

In this section, we shall analyse the growth of the macroeconomy, given the logistic growth of individual final goods (4.10).

The Basic Result

The aggregate value added or GDP of this economy is stochastic, but in what follows, we shall focus on its expected value and denote it by $Y(t)$. $Y(t)$ is simply the sum of production of all the final goods. Since profits in the final good sectors are zero by the assumption of perfect competition, the aggregate value added is equal to the value added (profit) produced by capital K in the intermediate good sector, $P_X X(t)$ which is equal to $\Sigma_k A X_k = \Sigma_k y_k = Y(t)$.

Figure 4.3 illustrates this model economy. Each sector once it emerged grows logistically. New sectors emerge stochastically, and the aggregate value added or GDP is simply the sum of outputs of all the then existing sectors.



Note: τ_i is the date of birth of the ith good/industry.

Figure 4.3 Saturation of demand and emergence of new final goods or industries

From (4.9) and (4.10), we know that the expected value of GDP of this economy is given by:

$$Y(t) = \sum_{N=1}^{\infty} \int_0^t \lambda N e^{-\lambda\tau}(1 - e^{-\lambda\tau})^{N-1} y_\tau(t) d\tau + \frac{\mu}{[\delta + (\mu - \delta)e^{-\mu t}]} \tag{4.11}$$

$$= \sum_{N=1}^{\infty} \int_0^t \lambda N e^{-\lambda\tau}(1-e^{-\lambda\tau})^{N-1} \frac{\mu}{[\delta + (\mu-\delta)e^{-\mu(t-\tau)}]} d\tau$$
$$+ \frac{\mu}{[\delta + (\mu-\delta)e^{-\mu t}]}.$$

The second term on the right-hand side is simply output of the 'first' sector at time t, $y_0(t)$. Using:

$$\lambda N e^{-\lambda\tau}(1-e^{-\lambda\tau})^{N-1} = \frac{d}{d\tau}(1-e^{-\lambda\tau})^N$$

and

$$\sum_{N=1}^{\infty}(1-e^{-\lambda\tau})^N = e^{\lambda\tau} - 1,$$

we obtain

$$Y(t) = \int_0^t \frac{\left[\frac{d}{d\tau}(e^{\lambda\tau}-1)\right]\mu}{[\delta + (\mu-\delta)e^{-\mu(t-\tau)}]} d\tau + \frac{\mu}{[\delta + (\mu-\delta)e^{-\mu t}]}$$

$$= \lambda \int_0^t \frac{e^{\lambda\tau}\mu}{[\delta + (\mu-\delta)e^{-\mu(t-\tau)}]} d\tau + \frac{\mu}{[\delta + (\mu-\delta)e^{-\mu t}]} \tag{4.12}$$

$$= \lambda \int_0^t \frac{e^{\lambda(t-u)}\mu}{[\delta + (\mu-\delta)e^{-\mu u}]} du + \frac{\mu}{[\delta + (\mu-\delta)e^{-\mu t}]}.$$

From (4.12), the growth rate of the expected value of GDP, g_t becomes:

$$g_t = \frac{\dot{Y}(t)}{Y(t)} = \lambda + \left[\frac{f(t)}{Y(t)}\right]\left[\frac{\dot{f}(t)}{f(t)}\right],$$

where $f(t)$ is the logistic equation:

$$f(t) = \frac{\mu}{[\delta + (\mu-\delta)e^{-\mu t}]}.$$

It is easy to show that g_t satisfies:

$$\dot{g}_t = (g_t - \lambda)[2(\mu-\delta)e^{-\mu t}f(t) - \mu - g_t] \tag{4.13}$$

with the initial value g_0.

$$g_0 = \frac{\dot{Y}(t)}{Y(t)}\Big|_{t=0} = \lambda + \mu - \delta.$$

Also, since $e^{-\mu t}f(t)$ approaches zero, we can establish that the growth rate of GDP asymptotically approaches λ.

$$\lim_{t\to\infty} g_t = \lim_{t\to\infty} \frac{\dot{Y}(t)}{Y(t)} = \lambda.$$

The growth rate of the economy is initially higher than λ by $\mu - \delta$, but it eventually goes down to λ. In this model, μ and δ have level effects while the ultimate growth rate is determined by λ. The exact time path depends, of course, on all the parameters μ, δ, and λ.

It is important to recognize that not only the steady-state growth, but also the out-of-steady-state growth is generated by the successive emergence of new products/industries. The growth of older industries keeps declining, while newer products/industries enjoy high growth. How high depends on μ and δ.

From the perspective of this model, it is easy to understand that historians have identified the 'leading' or 'key' industries in the process of economic growth. The best-known example would be perhaps Rostow (1960, pp. 261–2) who argues:

> The most cursory examination of the growth patterns of different economies, viewed against a background of general historical information, reveals two simple facts:
> 1. Growth-rates in the various sectors of the economy differ widely over any given period of time;
> 2. In some meaningful sense, over-all growth appears to be based, at certain periods, on the direct and indirect consequence of extremely rapid growth in certain particular key sectors.

Vigor of the leading sectors depends on μ and δ in the model. For the sake of illustration, we show a simulation result (Table 4.1 and Figure 4.4). In this example, we assume that λ, μ, and δ are 0.03, 0.12 and 0.02, respectively. Table 4.1 and Figure 4.4 show both the growth rate of GDP and the average growth rate defined as $\Sigma_{\tau=1}^{t} g_t/t$ for each period (year). For the first ten years, the growth rate of the economy is higher than 9 percent. In year 20, it is still 5.7 percent. By year 40 the growth rate has slowed to 3.2 percent which is close to the assumed asymptotic rate 3 percent. The average growth rate, of course, decelerates much more slowly than the growth rate itself. The average growth rate for the first 30 years, for example, is 7.5 percent, although the growth rate in the year 30 is 3.9 percent. This example demonstrates that depending on μ and δ, the economy can sustain a much higher growth rate than the equilibrium rate for a very long period. The transitionary dynamics are determined by the pattern of demand saturation.

Table 4.1 A simulation result ($\lambda = 0.03$, $\mu = 0.12$, $\delta = 0.02$)

Time	GDP	Growth rate (%)	Average growth rate (%)	Time	GDP	Growth rate (%)	Average growth rate (%)
0	1.00	–	–	51	18.82	3.0	5.8
1	1.14	12.8	12.8	52	19.40	3.0	5.7
2	1.28	12.3	12.5	53	20.00	3.0	5.7
3	1.45	11.8	12.3	54	20.62	3.0	5.6
4	1.62	11.4	12.1	55	21.25	3.0	5.6
5	1.81	11.0	11.9	56	21.90	3.0	5.5
6	2.01	10.6	11.6	57	22.57	3.0	5.5
7	2.23	10.2	11.4	58	23.26	3.0	5.4
8	2.46	9.8	11.2	59	23.98	3.0	5.4
9	2.70	9.4	11.0	60	24.71	3.0	5.3
10	2.95	9.0	10.8	61	25.46	3.0	5.3
11	3.22	8.7	10.6	62	26.24	3.0	5.3
12	3.50	8.3	10.4	63	27.04	3.0	5.2
13	3.79	7.9	10.2	64	27.87	3.0	5.2
14	4.09	7.6	10.1	65	28.72	3.0	5.2
15	4.39	7.2	9.9	66	29.59	3.0	5.1
16	4.71	6.9	9.7	67	30.50	3.0	5.1
17	5.03	6.6	9.5	68	31.43	3.0	5.1
18	5.35	6.3	9.3	69	32.38	3.0	5.0
19	5.68	6.0	9.1	70	33.37	3.0	5.0
20	6.01	5.7	9.0	71	34.39	3.0	5.0
21	6.35	5.4	8.8	72	35.44	3.0	5.0
22	6.69	5.2	8.6	73	36.52	3.0	4.9
23	7.03	5.0	8.5	74	37.63	3.0	4.9
24	7.37	4.8	8.3	75	38.78	3.0	4.9
25	7.72	4.6	8.2	76	39.96	3.0	4.9
26	8.07	4.4	8.0	77	41.17	3.0	4.8
27	8.42	4.2	7.9	78	42.43	3.0	4.8
28	8.77	4.1	7.8	79	43.72	3.0	4.8
29	9.13	4.0	7.6	80	45.05	3.0	4.8
30	9.48	3.9	7.5	81	46.42	3.0	4.7
31	9.85	3.8	7.4	82	47.84	3.0	4.7
32	10.21	3.7	7.3	83	49.30	3.0	4.7
33	10.59	3.6	7.2	84	50.80	3.0	4.7
34	10.96	3.5	7.0	85	52.34	3.0	4.7
35	11.35	3.4	6.9	86	53.94	3.0	4.6
36	11.74	3.4	6.8	87	55.58	3.0	4.6
37	12.14	3.3	6.7	88	57.27	3.0	4.6
38	12.54	3.3	6.7	89	59.02	3.0	4.6
39	12.96	3.3	6.6	90	60.82	3.0	4.6
40	13.38	3.2	6.5	91	62.67	3.0	4.5
41	13.81	3.2	6.4	92	64.58	3.0	4.5
42	14.26	3.2	6.3	93	66.54	3.0	4.5
43	14.71	3.1	6.3	94	68.57	3.0	4.5
44	15.18	3.1	6.2	95	70.66	3.0	4.5
45	15.66	3.1	6.1	96	72.81	3.0	4.5
46	16.15	3.1	6.0	97	75.03	3.0	4.5
47	16.66	3.1	6.0	98	77.31	3.0	4.4
48	17.18	3.1	5.9	99	79.67	3.0	4.4
49	17.71	3.1	5.9	100	82.09	3.0	4.4
50	18.26	3.1	5.8				

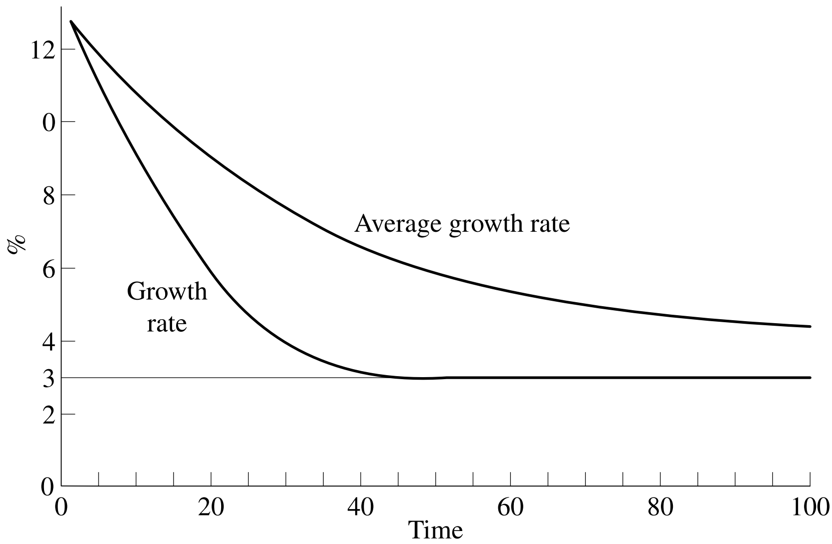


Figure 4.4 A simulation result (λ=0.03, μ=0.12, δ=0.02)

Everyone knows that no economy grows at 10 percent indefinitely. Some economies, however, actually experienced 10 percent growth for a decade, and this decade-long high growth is often crucial for their growth experiences. Japan, for example, experienced 10 percent growth for a decade and a half, from 1955 to 1970. As of 1955, almost half of the working population in Japan was in agriculture. The era of high economic growth had transformed a semi-traditional economy into a modern industrial nation. We cannot dismiss 'out of steady state' merely as transitory, but must attach equal importance to it as to the steady state.

The out-of-steady-state growth path illustrated in Figure 4.4 is qualitatively similar to that obtained in the old Solow (1956) model and also in the more recent R&D-based growth models such as Jones (1995), Segerstrom (1998) and Young (1998). Namely, the growth rate decelerates over time. The mechanism is fundamentally different, however. In the Solow model, diminishing returns to capital in ordinary production is the factor which brings about slower growth. In the R&D-based models, diminishing returns in R&D leads to slower growth. In contrast, in the present model the deceleration of the out-of-steady-state growth rate comes from saturation of demand. To be specific, as equation (13) shows, the out-of-steady state growth path depends on μ and δ, which determine how vigorous growth of demand is, or, conversely, how soon demand reaches its saturation. When δ is very small and, therefore, the 'ceiling' of demand is very high, the pace of deceleration can be slow. This out-of-steady-state growth is basically constrained by demand. In this sense, the present model follows the long tradition of post-Keynesian literature mentioned in Section 1.

An Extension: The Non-Poisson 'Polya Urn' Model

In the model above, we assumed that a new good or sector emerged as a branch off in the production of N existing goods and that it followed the Poisson birth process with the parameter λN. Under this assumption, the long-run growth rate is sustained by the rate of innovation λ.

In the standard R&D-based growth models, Jones (1995), Jones and Williams (1998), Segerstrom (1998) and Young (1998), among others, showed that if R&D was subject to diminishing returns due, for example, to congestion in research, the finishing-out effect and increasing difficulty, then the growth rate decelerated to zero in the steady state unless some exogenous factor such as population growth sustained it.

We shall examine a similar problem in our model. In the existing literature mentioned above all of which use Poisson model, the Poisson parameter, which is λ in our model, is the success probability of R&D and, therefore, the birth rate is naturally taken as a function of the stock of R&D, which corresponds to N in our model. The deceleration of the long-run growth rate occurs when the birth rate $\lambda = f(N)$ shows diminishing returns, namely when $f(N) / N$ is a decreasing function of N.

In our basic model, the birth of a new goods/sector is a branch off from the production of N existing goods. Therefore, we assumed that the birth rate was λN. However, we model the emergence of wholly new goods/industries rather than close substitutes of the existing goods. It is not directly linked to R&D but is strongly conditioned on the advancement of basic scientific knowledge. One might like to assume, therefore, that if opportunities for innovations narrow over time, the probability of the emergence of new goods/sectors would be a decreasing function of time rather than N. Specifically, in place of the Poisson distribution, we assume that the probability that a new good or sector emerges at τ, p_τ is:

$$p_\tau = \frac{\omega}{\omega + \tau} \quad (\omega > 0, \tau = 1, 2, \ldots).$$

This probability decreases in τ, and declines asymptotically to zero. This kind of model, often called 'Polya-like urns', is extensively used in population genetics (see, for example, Hoppe 1984). The existing literature in economics all relies on the Poisson distribution and, to our knowledge, the non-Poisson model is new.

Now, we assume that a new good is invented 'exogenously' with p_τ rather than as a branch off from the existing goods, namely, that p_τ is independent of the number of existing goods. In this case, when we denote the probability that there are N goods at τ by $Q(N,\tau)$ as we did previously, then $Q(N,\tau)$ satisfies:

$$Q(N,\tau+1) = (1-p_\tau)Q(N,\tau) + p_\tau Q(N-1,\tau)$$
$$= \left(\frac{\tau}{\omega+\tau}\right) Q(N,\tau) + \left(\frac{\omega}{\omega+\tau}\right) Q(N-1,\tau), \text{ for } \tau = 1, 2, \ldots,$$

with the following boundary conditions:

$$Q(1,\tau) = \left(\frac{1}{\omega+1}\right)\left(\frac{2}{\omega+2}\right)\cdots\left(\frac{\tau-1}{\omega+\tau-1}\right)$$

and

$$Q(\tau,\tau) = \frac{\omega^\tau}{\omega(\omega+1)(\omega+2)\cdots(\omega+\tau-1)} = \frac{\omega^\tau}{[\omega]^\tau},$$

where $[\omega]^\tau$ is defined by the equation.

The solution of this equation is:

$$Q(k,\tau) = \frac{c(k,\tau)\omega^k}{[\omega]^\tau},$$

where $c(k, \tau)$ is the absolute value of the Sterling number of the first kind (see Abramovitz and Stegun 1968, p. 825 or Aoki 1996, 1997, p. 279). Using the generating function:

$$[x]^k = \sum_{j=0}^{k}, c(k,j)x^j,$$

we obtain the expected value of GDP, $Y(t)$ as:

$$Y(t) = \sum_{\ell=1}^{t}\sum_{j=1}^{\ell} \frac{c(\ell-1, j-1)\omega^{j-1}}{[\omega]^{\ell-1}}\left(\frac{\omega}{\omega+\ell}\right) y(t-l)$$
$$= \sum_{\ell=1}^{t}\left(\frac{\omega}{\omega+\ell}\right) y(t-\ell).$$

Here $y(t-\ell)$ is the production of the final good which emerged at time ℓ. Note that $y(t-\ell)$ follows the logistic curve, and therefore, that its growth rate eventually declines to zero.

For simplicity, take ω as an integer. Then we have:

$$\sum_{\ell=1}^{t}\left(\frac{\omega}{\omega+\ell}\right)=\omega\left(\sum_{m=1}^{\omega+t}\frac{1}{m}-\sum_{m=1}^{\omega}\frac{1}{m}\right)\cong\log\left(\frac{\omega+t}{\omega}\right).$$

Therefore we have shown that in the present case, GDP grows, drawing the logarithmic curve:

$$Y(t)\sim\log(t+\omega).$$

The growth rate of the economy is $1/t+1[t\log(t+\omega)]$, and goes asymptotically down to zero.

Thus, the result similar to Jones (1995), Segerstrom (1998) and Young (1998) holds in our model; if the opportunities for innovations diminish over time, the long-run growth is not sustained. Note that in the existing literature, λ is a (possibly) decreasing function of N whereas in the present analysis, λ is independent of N and is a decreasing function of time.

4. FOUNDATIONS FOR THE LOGISTIC GROWTH OF DEMAND

Having found the growth rate of GDP, we next turn to the general equilibrium of this model. We have already explained production of both final goods and an intermediate good. In what follows, we first consider the firm behavior which determines investment in the intermediate good sector, and then the consumer behavior which determines consumption/saving. The consumer behavior must be consistent with growth of income or GDP (equation (4.13)) and the logistic growth of demand for an individual final good. Also saving must be equal to investment; final goods are not only consumed but also used for investment.

In this model, consumption leads to the logistic growth of individual final goods. We suggest two different models, one the standard Ramsey model with the representative consumer and the other with diffusion of final goods among different households. Since the model of investment is common, we begin with the firm's investment.

The Firm's Investment Decisions

The intermediate good is produced by the representative firm using capital K (see equation (4.6)). The firm is constrained by demand, and the capacity utilization rate a is variable. Capital accumulates so as to maximize the value of this firm (industry). Profit of this firm, which stems from selling intermediate goods to firms producing final goods, is P_xX.

Gross investment I requires finished goods as an input. For simplicity, we assume that final goods are perfect substitutes for the purpose of increasing K. I consists of two parts: the net investment inclusive of the standard adjustment costs; and the depreciation, which depends positively on the capacity utilization rate (a in equation (4.6)) of the existing capital stock K. Specifically, we assume:

$$I = \varphi(z)K + d(a)K, \tag{4.14}$$

with $z = K/K$.

The variable $\varphi(z)$ satisfies:

$$\varphi' = (z) > 0 \text{ for } z = \dot{K}/k > 0,\ \varphi'(z) < 0 \text{ for } z < 0,\ \varphi''(z) > 0 \text{ for any } z,$$
$$\varphi(0) = 0, \text{ and } \varphi'(0) = 1. \tag{4.15}$$

We assume the convex adjustment cost for negative z to rule out disaccumulation of capital. Depreciation $d(a)$ is also a convex function of the capacity utilization rate (Smith 1969):

$$d'(a) > 0,\ d''(a) > 0, \text{ and } d(0) = 0. \tag{4.16}$$

The value of this firm S is then given by:

$$S_t = \int_t^\infty (P_X X_\tau - I_\tau)\exp\left(-\int_t^\tau \rho_u du\right)d\tau. \tag{4.17}$$

$$= \int_t^\infty [P_X X_\tau - \varphi(Z_\tau)K_\tau - d(a_\tau)K_\tau]\exp\left(-\int_t^\tau \rho_u du\right)d\tau.$$

We note that S_t satisfies:

$$\rho_t = \frac{\dot{S}_t}{S_t} + \frac{P_X X_t - I_t}{S_t} \tag{4.18}$$

and observe that ρ is the rate of return on 'stock' of this firm or the interest rate. The stock of this firm is owned by the consumer (or consumers). This will be explained shortly.

The firm is constrained by demand X. Since capital stock K is also given, the capacity utilization rate $a = X/K$ is given at each moment in time. The firm maximizes its present value with respect to investment. If the pace of capital accumulation is short of the growth rate of sales, the capacity utilization rate rises, and so does depreciation. On the other hand, when the firm raises the rate of capital accumulation, it incurs higher adjustment costs. The firm must balance the two so as to maximize its present value.

This decision depends crucially on the growth rate of demand $g = X/X$ facing the firm.

A change in the capacity utilization rate $a = X/K$ must by definition satisfy the following equation:

$$\dot{a}_t = (g_t - z_t)a_t \tag{4.19}$$

It can be shown that the optimal capital accumulation z and capacity utilization rate a must satisfy (4.19) and (4.20):

$$\dot{z} = \left[\frac{1}{\varphi''(z)}\right][(\rho - z)\varphi'(z) + \varphi(z) - r(a)] \tag{4.20}$$

where:

$$r(a) = d'(a)a - d(a), \quad r'(a) > 0 \tag{4.21}$$

where:
$r(a)$ measures a marginal increase in the firm's value when the rate of capital accumulation z rises by way of lowering the capacity utilization, and accordingly depreciation. Therefore, $r(a)$ is the profit rate on capital K in this model. Given capital stock K, the profit rate r is an increasing function of the firm's demand X. The rule for the optimal capital accumulation (4.20) is basically equivalent to the one in Uzawa (1969).

This optimal path can be most clearly seen with the help of Figure 4.5 when g is constant as in the steady state. The capital accumulation z approaches the growth rate of sales g from above when the initial capacity utilization is higher than its equilibrium level a^* which satisfies $r(a^*) = (\rho - g)\varphi'(g) + \varphi(g)$ and vice versa. When g becomes higher, so does the long-run profit $r(a^*)$, and higher $r(a^*)$ induces higher capital accumulation. The firm's prospects for demand basically determine investment. Investment also depends negatively on the interest rate ρ.

So much for the investment of the demand-constrained firm. In what follows, we shall consider the consumer behavior which determines consumption/saving. We shall present two alternative stories: the Ramsey model, and a model of diffusion of consumption goods among different households.

The Consumer's Consumption/Saving Decisions

The Ramsey Model

In the neoclassical approach, the queen of the economy is the consumer. Demand for final goods must, therefore, be consistent with the consumer's utility maximization. In what follows, we demonstrate that demand for final

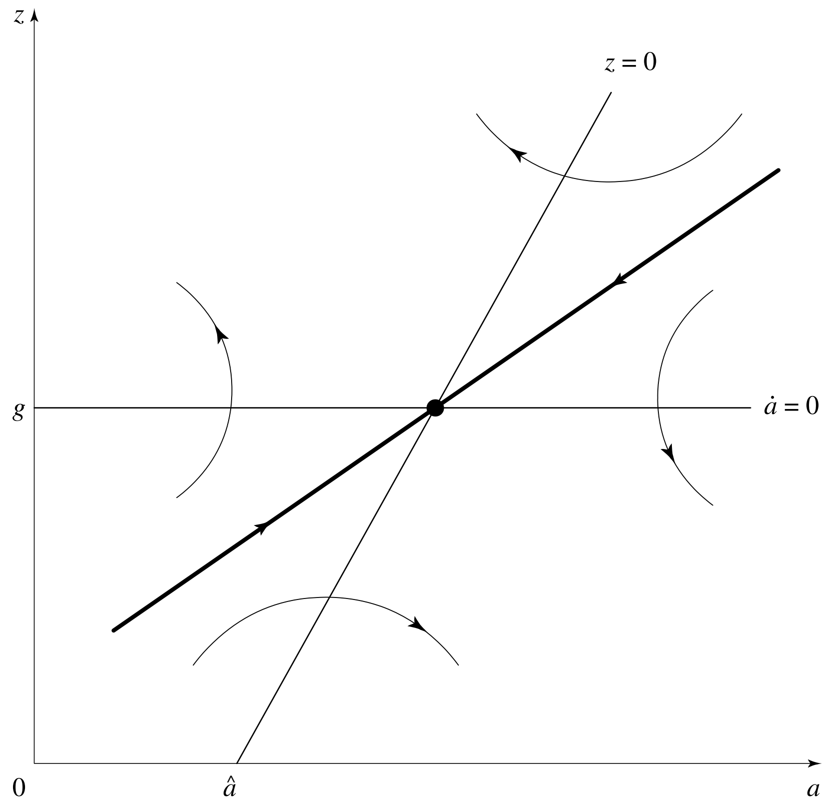


Figure 4.5 *The determination of capital accumulation* z *and capacity utilization rate* a *(*$\hat{a}$ *is defined by* $\rho = d'(\hat{a} - d(\hat{a}))$*)*

goods which obeys the logistic equation is in fact consistent with the intertemporal utility maximization of the Ramsey consumer with a particular utility function.

For convenience, we consider the representative consumer's utility maximization at time 0. At time 0, there is only one final good as is assumed in Section 3. This assumption is made just for simplicity. The assumption that there are n goods ($n>0$) at time 0 merely deprives our presentation of its simplicity without giving us any additional insight.

Starting with one final good at time 0, new goods keep emerging. The probability that there are N goods at time t and the $N+$ good emerges during t and $t+\Delta t$ is given in Equation (4.13). Thus as of time 0, the consumer faces uncertainty concerning the timing of the emergence of new goods. We assume that the consumer maximizes the expected utility:

$$U=\int_0^{\infty}\left(\int_0^{t}\sum_{N=1}^{\infty}\{\lambda N e^{-\lambda\tau}(1-e^{-\lambda\tau})^{N-1}u_{\tau}^{t}[C_{N+1}(t)]\}d\tau+u_0^t[C_1(t)]\right)e^{-\theta t}dt, \quad (4.22)$$

where θ is the subjective discount rate and $C_j(t)$ is the consumption of the jth good. In (4.22), the expected value is taken with respect to the probability of the emergence of new goods. A similar assumption is made in Aghion and Howitt (1992).

To obtain the logistic demand function, we assume that the utility coming from consumption of a certain final good at time t depends not only on t but also on τ ($\tau < t$), the time when this final good emerged. To be specific, we assume that the utility function $u_{\tau}^{t}[C_N(t)]$ is common for all the $C_N(t)$ (N = 1, 2, . . .) and is:

$$u_{\tau}^{t}[C_N(t)]=\left\{\frac{\mu}{[\delta+(\mu-\delta)e^{-\mu(t-\tau)}]}\right\}\log\,[C_N(t)]. \quad (4.23)$$

The logistic growth of demand (4.14) characterized by two parameters μ and δ translates itself into the utility function (4.23). It is actually more accurate to say that μ and δ which characterize the time-dependent utility function lead us to the logistic growth of demand for an individual final good. The logistic part of utility function (4.23) implies that the (marginal) utility coming from consumption of a particular final good depends crucially on how much time has passed since this final good first emerged. Though it monotonically increases over time, its growth rate eventually decelerates and is bound to approach zero.

The consumer owns the stock (capital) of the intermediate good industry, S_t which earns the rate of return ρ. Thus his/her budget constraint is:

$$\dot{S}_t=\rho_t S_t-\sum_{i=1}^{\infty}C_i(t). \quad (4.24)$$

The consumer maximizes (4.22) subject to (4.24) and S_0. Introducing the costate variable (shadow price of capital stock) $\nu(t)e^{-\theta}$, we obtain the necessary conditions for optimality as follows:

$$C_1(t)=\left[\frac{1}{\nu(t)}\right]\left[\frac{\mu}{\delta+(\mu-\delta)e^{-\mu t}}\right] \quad (4.25)$$

$$C_{N+1}(t)=\int_0^t \lambda N e^{-\lambda\tau}(1-e^{-\lambda\tau})^{N-1}\left[\frac{1}{\nu(t)}\right]\left[\frac{\mu}{\delta+(\mu-\delta)e^{-\mu(t-\tau)}}\right]d\tau \quad (4.26)$$

$$\frac{\dot{\nu}(t)}{\nu(t)} = \theta - \rho_t \tag{4.27}$$

and

$$\lim_{t\to\infty} \nu(t)e^{-\theta t}S(t) = 0. \tag{4.28}$$

Since $S(t)$ grows asymptotically at the rate of $g(t)$ and $\nu(t)$ satisfies (4.27), the transversality condition (4.28) is equivalent to:

$$\lim_{t\to\infty} \exp\left[-\int_0^t (\rho_\tau - g_\tau)d\tau\right] = 0. \tag{4.29}$$

Condition (4.29) is satisfied when the optimal solution (4.20) exists for investment decisions.

From (4.25) and (4.26), we obtain:

$$C(t) = \sum_{j=1}^{\infty} C_j(t) = \left[\frac{1}{\nu(t)}\right]\left\{\frac{\mu}{[\delta + (\mu - \delta)e^{-\mu t}]}\right. \tag{4.30}$$

$$\left. + \sum_{N=1}^{\infty} \int_0^{\infty} \lambda N e^{-\lambda\tau}(1 - e^{-\lambda\tau})^{N-1} \frac{\mu}{(\delta + (\mu - \delta)e^{-\mu(t-\tau)})}\right] d\tau \Bigg\}$$

for total consumption $C(t)$ at time t. Due to (4.11), we can rewrite (4.30) as:

$$C(t) = \frac{Y(t)}{\nu(t)}. \tag{4.31}$$

$1/\nu(t)$ is simply the average propensity to consume.

Given (4.27), equation (4.31) is equivalent to:

$$\theta - g_t + \frac{\dot{C}(t)}{C(t)} = \rho_t. \tag{4.32}$$

Equation (4.32) is the Euler equation or the Keynes/Ramsey rule, which requires that for optimality, the marginal rate of substitution defined by the left-hand side of equation (4.32) must be equal to the interest rate ρ. The optimal saving decisions satisfy (4.32).

As we have already seen, the optimal investment decisions satisfy (4.20). Note that both the optimal saving and investment decisions (4.20) and (4.32) depend on the time paths of the growth rate of Y, g_t and the interest rate ρ_t. The interest rate ρ_t, the growth rate of capital z_t, and the capacity utilization a_t are simultaneously determined by (4.19), (4.20) and (4.32) in such a way that g_t satisfies (4.13).

The goods market equilibrium, namely:

$$Y = C + I = \sum_i C_i + \phi(z)K + d(a)K \tag{4.33}$$

is, as usual, assured by the appropriate change in the interest rate. Formally, the time path of this equilibrium interest rate ρ can be found by considering the 'command economy' corresponding to the market economy. Maximize the consumer's utility (4.22) under the constraint (4.19) and (4.33). The time path of Y_t is given by (4.13) and sets a constraint for this problem. Call the Lagrange multiplier for the goods market equilibrium constraint (4.33) $\omega_t e^{-\theta t}$. Then $\rho_t = \theta - (\dot{\omega}/\omega)$ is the equilibrium interest rate.

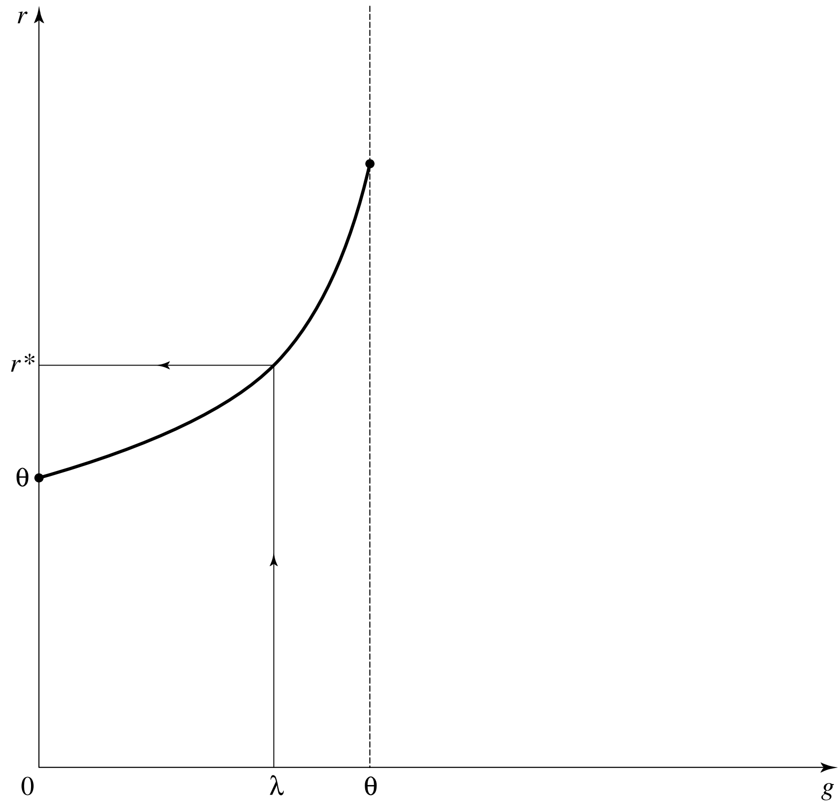
In this model, the capacity utilization rate a and accordingly the profit rate $r(a) = d'(a)a - d(a)$ are endogenously determined. The higher the growth rate determined by δ, μ and λ, the higher the profit rate r. This can be seen most clearly for the steady state. In the steady state, $\dot{C}/C$ is g, and the interest rate ρ becomes equal to the consumer's discount rate θ. The relation between the growth rate g and the profit rate r in this case is shown in Figure 4.6. When the strength of innovations to create new sectors/goods λ gets higher, the steady-state growth rate becomes higher. Higher investment is induced by higher profit which is in turn generated by the higher capacity utilization rate.

To equilibriate the goods market, higher investment (or a 'shift up' of the investment function) generated by high growth of demand brings about a high interest rate. The high interest rate in turn makes the consumer find high growth of consumption desirable (the Euler equation). This accordingly generates high saving which must be equal to investment in equilibrium. The ultimate factor to determine the growth rate is the 'vigor of demand' characterized by three parameters μ, δ and λ.

In the standard growth models such as Romer (1990), Grossman and Helpman (1991), Aghion and Howitt (1992) and Jones (1995), the rate of innovation raises the long-run growth rate of the economy by increasing TFP. In the present model, the rate of innovation λ raises the growth rate of the economy by increasing aggregate demand. Technical progress elicits investment of the demand-constrained firm both by increasing the number of goods produced over time and, transitionally, by reducing the average age of products in the market, favoring products on faster growth segments of their demand life cycles.

Diffusion of final goods among different households

Ramsey model is the most standard approach in macroeconomics. However, in many economies, for many periods in history, a declining growth in demand for a particular product has been very closely related to diffusion of the product among *different* households. Some households own the product while others do not. This is particularly true for such consumer durables as televisions, refrigerators, cars and personal computers. For these consumer durables, it makes more sense to analyse their growth



Note: The relation between g and r is given by $r = (\theta - g)\varphi'(g) + \varphi(g)$. Thus $r^* = (\theta - \lambda)\varphi'(\lambda) + \varphi(\lambda)$.

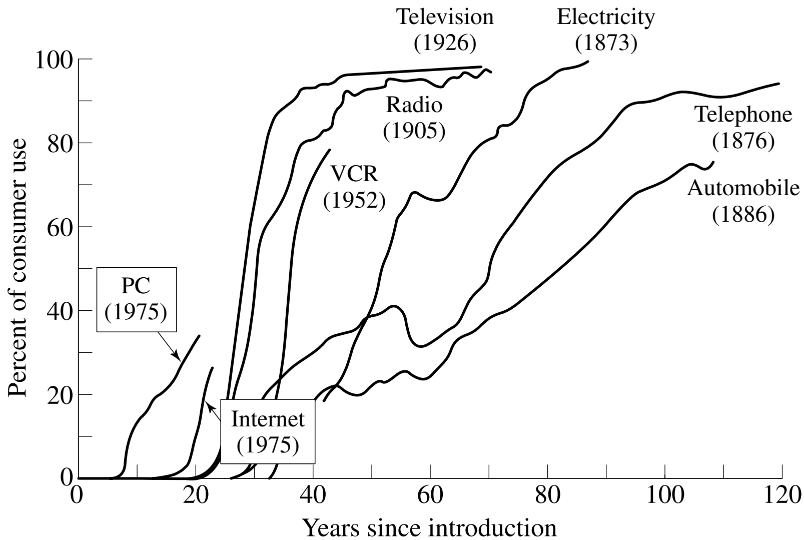
Figure 4.6 The determination of the profit rate r *corresponding to the steady-state growth rate* λ

in a model with different households than in a model with the representative consumer. Diffusion of consumer goods among households, in fact, plays an important role in the logistic growth.[4] Bill Gates notes this fact (Figure 4.7; Gates 1999, p. 118). In this section, we consider such a model.

Suppose that there are M households in the economy. Without loss of generality, we can assume that M is equal to μ/δ. Households are indexed by i ($i = 1, 2, \ldots, M = \mu/\delta$). We define $f_{iN}(t)$ functions:

$f_{iN}(t) = 1$, if household i purchases the Nth product at time t.

$f_{iN}(t) = 0$, if household i does not purchase the Nth product at time t.



Source: Gates (1999).

Figure 4.7 Diffusion of consumer durables

We assume that the number of households which consume the Nth product at time t, $m_N(t)$ follows a birth and death process with the birth rate μ and the death rate δm_N. Note the following relation:

$$\sum_{i=1}^{M} f_{iN}(t) = m_N(t). \tag{4.34}$$

This process leads us to the logistic equation for the expected value of $m_N(t)$, $\hat{m}_N(t)$. Thus if the Nth product emerged at τ, $\hat{m}_N(t)$ satisfies the following equation:

$$\hat{m}_N(t) = \frac{\mu}{[\delta + (\mu - \delta)e^{-\mu(t-\tau)}]} \quad \text{for each } N. \tag{4.35}$$

The (expected) diffusion rate or the percentage of households which consume the Nth product is $\hat{m}_N(t)/M$.

For simplicity we assume that a household purchases $1-s$ units of any final product if it consumes this product, s is the saving rate. As in Solow (1956), the present analysis abstracts from the determination of s. The saving rate is assumed to be common for all the households ($i = 1, 2, \ldots, M$), and depends positively on the interest rate ρ_t and on time t. Note that the consumption of a final good by each household is constant. This assumption seems to hold, as an approximation, for many consumer durables.

Then, with the expected income of household i, $I_i(t)$, the budget constraint for household i at time t becomes:

$$I_i(t) = \sum_{N=1}^{\infty} \int_0^t \lambda N e^{-\lambda\tau}(1-e^{-\lambda\tau})^{N-1}(1-s)f_{iN+1}(t)d\tau + (1-s)f_{i1}(t) + sI_i(t)$$

$$= \sum_{N=1}^{\infty} \int_0^t \lambda N e^{-\lambda\tau}(1-e^{-\lambda\tau})^{N-1} f_{iN+1}(t)d\tau + f_{i1}(t). \tag{4.36}$$

Due to (4.34) and (4.35), incomes of all the households (4.36) sum up to GDP, $Y(t)$:

$$\sum_{i=1}^{M} I_i(t) = \sum_{N=1}^{\infty} \int_0^t \lambda N e^{-\lambda\tau}(1-e^{-\lambda\tau})^{N-1} \sum_{i=1}^{M} f_{iN+1}(t)d\tau + \sum_{i=1}^{M} f_{i1}(t)$$

$$= \sum_{N=1}^{\infty} \int_0^t \lambda N e^{-\lambda\tau}(1-e^{-\lambda\tau})^{N-1} \frac{\mu}{[\delta + (\mu-\delta)e^{-\mu(t-\tau)}]} d\tau$$

$$+ \frac{\mu}{[\delta + (\mu-\delta)e^{-\mu t}]}$$

$$= Y(t).$$

The goods market equilibrium is then:

$$s(\rho_t) = \frac{\varphi(z_t)K_t + d(a_t)K_t}{Y_t} = \frac{\varphi(z_t) + d(a_t)}{a_t}. \tag{4.37}$$

Just as in the Ramsey model, the interest rate ρ, the growth rate of capital z_t, and the capacity utilization rate a_t are simultaneously determined by (4.19), (4.20), and (4.37) in such a way that g_t satisfies equation (4.13) in Section 3. In the steady state, $s(\rho^*) = [\varphi(\lambda) + d(a)^*]/a^*$ and $\varphi'(\lambda) = [r(a^*) - \varphi(\lambda)]/(\rho^* - \lambda$ hold. The steady-state growth rate is λ. The steady-state values of the interest rate ρ^* and capacity utilization rate a^* are different from those of the Ramsey model. The effects of an increase in λ on the equilibrium interest rate and the capacity utilization rate are, however, qualitatively the same as in the Ramsey model provided that the interest elasticity of the saving rate is high enough.

In the Ramsey model with the representative consumer, the S-shaped growth of each product derives directly from the assumption (4.23) that the marginal utility coming from consumption of a particular final good depends on how much time has passed since the good first emerged. In the second model, a household is assumed to purchase a given unit of each product. The S-shaped growth derives from diffusion of each product among different households. In both cases, growth of demand induces capital accumulation by the demand-constrained firm, and thereby economic growth. Growth, on the other hand, creates higher income which enables more households able to purchase goods. Equation (4.36) defines income distribution which generates diffusion of final goods among households. Because the amount of a final good which each household purchases is bounded, growth of production of an individual good necessarily decelerates parallel to diffusion of the good among households. Creation of wholly new goods/sectors is the ultimate factor to sustain growth in such an economy.

5. CONCLUDING REMARKS

In the standard literature, the fundamental factor restraining economic growth is diminishing returns to capital in production or R&D technology. We presented a model in which the factor restraining growth was saturation of demand. Our analysis began with a common observation that for individual products/industries, there was a history of logistic development with initial acceleration and eventual retardation of growth. Taking this as a 'stylized fact', we presented a formal model of growth consistent with this important 'fact'.

This model provides new perspectives to several important problems addressed by the economics of growth. The first new perspective pertains to out-of-steady-state dynamics. Despite the controversy surrounding conditional convergence, for many economies, and for advanced economies in particular, we do observe the eventual deceleration of growth rates. In the Solow model, diminishing returns to capital in production is the factor to bring about slower growth. In the R&D based models, it is diminishing returns in R&D. Thus, in both approaches, diminishing returns in technology is the factor that brings about a slowdown of economic growth. In contrast, in our model, saturation of demand (μ and δ) is the factor that leads the economy to slower growth.

The second new perspective relates to the nature of technical progress or innovations. In the standard analysis, technical progress brings about higher value added given the same level of inputs. It is basically equivalent to an 'upward shift' of the production function. The so-called 'product variety', 'quality ladder' or 'creative destruction' models, such as Grossman and Helpman (1991) and Caballero and Jaffe (1993), using the Dixit/Stiglitz production/utility function, successfully endogenize this kind of technical progress. Empirically, technical progress has been measured by growth accounting as TFP.

In the model presented here, the aggregate production function is $Y = AK$. Since A is constant, there is no TFP growth; For the economy to grow, capital must accumulate. However, saturation of demand constrains capital accumulation and leads the economy to deceleration of growth. Innovation or technical progress in this model creates a major new product or industry which commands high growth of demand and thereby elicits capital accumulation and so sustains economic growth. In his famous book, Schumpeter (1934) distinguishes five types of innovation: (i) the introduction of a new good, (ii) the introduction of a new production method, (iii) the opening of a new market, (iv) the conquest of a new source of supply of raw materials, and (v) the new organization of industry. His first and third types of innovation as an engine for growth seem to be most naturally interpreted in terms of the kind of model presented here.

The distinction between the conventional TFP and demand-creating technical progress is not only theoretically important but is also empirically relevant. Young (1995), for example, in his careful study of growth accounting, demonstrates that TFP growth in the newly industrializing countries (NICs) of East Asia, that is, Hong Kong, Singapore, South Korea and Taiwan, is not so extraordinarily high, but actually comparable to that in other countries. The very high average growth rate (8 to 10 percent) of East Asian NICs for such a long period as 25 years must, therefore, be explained by extraordinary injection of capital and labor rather than extraordinary TFP growth.

Still the basic question remains: why did the economy grow so fast in these countries? The analysis of this chapter suggests that in these countries, new sectors which command high growth of demand emerged vigorously (high μ/δ and λ). TFP growth may not be extraordinarily high, but it does not necessarily mean there is an absence of demand-creating innovations. Young, in fact, reports that in East Asian NICs the industrial structure has drastically changed. A change in the industrial structure in the course of economic growth is likely to reflect demand-creating innovations which are conceptually different from TFP. To the extent that exports have commanded high growth (that is, high μ/δ), we can easily understand that the high growth of East Asian NICs often looked to be export led. Nelson and Pack (1999) present a similar argument, emphasizing the importance of productive assimilation for the success of Asian NICs. Productive assimilation is related to λ in our model.[5]

To repeat, this chapter shows that technical progress has a different aspect than that represented by conventional TFP, namely demand creation, and that it is demand creation that sustains economic growth. This may shed light not only on 'the Asian miracle' but also on the high growth of the American economy during the 1990s where despite remarkable progress in information technology (IT), TFP growth was not so high as one might expect at least until the late 1990s (Gordon 1999). Advancement of IT may not only increase productivity on the supply side, but also create new markets and mediate demand-led growth.

Our model also provides different policy implications from the standard R&D-based growth models. The standard literature amply demonstrates the importance of R&D. However, these models ignore differences in the age of products/industries. In our model, R&D in the mature product/industry does not promote economic growth *not* because efficiency in R&D diminished but primarily because demand approached saturation. Any policy to promote growth must seriously take into account the age of the product or the prospect of demand.

The ultimate factor in sustaining growth λ which creates wholly new products/industries would depend not only on profit-motivated R&D, but also heavily on basic scientific research. In fact, λ is not necessarily confined to supply-side factors. Growth and saturation of demand often parallel diffusion among different households, as we discussed in the second model in Section 4. Thus appropriate income distribution policy which triggers diffusion of major products can be taken as a demand-creating innovation. This point may be important for the growth of developing countries. It was certainly important in the post-war Japanese growth during the 1950s and 1960s.[6]

The importance of demand as a determinant of aggregate economic activity has been demonstrated in a variety of contexts by such authors as Diamond (1982) and Murphy et al. (1989). In the Diamond model, for example, an externality causes coordination failures in search equilibrium. In the Murphy et al. model of economic development, simultaneous industrialization of many sectors of the economy leads it from a 'bad' to a 'good' equilibrium through aggregate demand spillovers. One way or another, demand failures emphasized by Keynes (1936) play the key role to generate multiple equilibria in these models. Our model suggests that demand saturation plays a similar role in the context of economic growth. As a model of demand-constrained growth, it follows the long line of post-Keynesian literature mentioned in the introduction.

Solow (1997) emphasizes the importance of 'the medium-run' analysis as a challenge to modern macroeconomics:

> One major weakness in the core of macroeconomics as I have represented it is the lack of real coupling between the short-run picture and the long-run picture. Since the long-run and the short-run merge into one another, one feels they cannot be completely independent. There are some obvious, perfunctory connections: every year's realized investment gets incorporated in the long-run model. That is obvious. A more interesting question is whether a major episode

> in the growth of potential output can be driven from the demand side. (Solow 1997, pp. 231–2)

In short, the integration of the Keynes principle of effective demand for the short run and growth theory for the long run remains a central theme in macroeconomics. Our model may provide a constructive step toward solving this problem.

Finally, to make our analysis tractable, we inevitably made an unrealistic assumption that μ, δ, and λ were constant. We hope that the assumption is justified for the purpose of studying economic growth. However, in the short/medium run, μ, δ and λ would all fluctuate. Giving the μ, δ and λ shocks to the model economy, as is done in the standard real business cycle literature, one would be able to generate fluctuations of the growth rate. Such simulation exercises with 'demand shocks' might generate a more realistic explanation of short-run fluctuations than those based on TFP shocks.

APPENDIX 4A

Equation (4.7) in the text can be solved in the following way. First we define the generating function $G(z,t)$ as:

$$G(z,t) = \sum_{n=N_0}^{\infty} z^n Q(n,t).$$

Multiplying (4.7) by z^n, and taking its sum over $n = N_0, N_0 + 1, \ldots$, we obtain the partial differential equation:

$$\frac{\partial G(z,t)}{\partial t} = \lambda z(z-1)\frac{\partial (G,t)}{\partial z} \tag{4A.1}$$

with the initial condition:

$$G(z,0) = z^{N_0}. \tag{4A.2}$$

To solve this partial differential equation, we introduce the artificial variable, s, for which the following ordinary differential equations hold:

$$\frac{dt}{ds} = 1 \tag{4A.3}$$

$$\frac{dz}{ds} = -\lambda z(z-1). \tag{4A.4}$$

With the initial condition $(s, t) = (0, 0)$, (4A.3) can be solved immediately to give $t = s$.

Similarly, with the initial condition $(s, z) = (0, m)$, (4A.4) can be solved to give:

$$\lambda s = \log\left(\frac{z}{z-1}\right)\left(\frac{m-1}{m}\right). \tag{4A.5}$$

Since $s = t$, from (4A.5) we obtain:

$$m = \frac{e^{-\lambda t} z}{[1-(1-e^{-\lambda t})z]}. \tag{4A.6}$$

On the other hand, from (4A.1), (4A.3) and (4A.4) we know that $G[z(s)]$ $t(s))$ satisfies:

$$\frac{dG}{ds} = \frac{\partial G}{\partial z}\cdot\frac{dz}{ds} + \frac{\partial G}{\partial t}\cdot\frac{dt}{ds} = -\lambda z(z-1)\frac{\partial G}{\partial z} + \frac{\partial G}{\partial t} = 0,$$

and therefore that G as a function of s is constant. Since z is m when s is zero, from (4A.2) we find that this constant is m^{N_0}. Using (4A.6), we see that $G(z, t)$ is:

$$G(z,t) = \frac{e^{-\lambda N_0 t} z^{N_0}}{[1-(1-e^{-\lambda t})z]^{N_0}}. \tag{4A.7}$$

The denominator of (4A.7) can be expanded as:

$$\frac{1}{[1-(1-e^{-\lambda t})z]^{N_0}} = \sum_{\ell=0}^{\infty}\binom{-N_0}{\ell}(-1)^l(1-e^{-\lambda t})^l z^l$$
$$= \sum_{\ell=0}^{\infty}\left(\frac{N_0+\ell-1}{\ell}\right)(1-e^{-\lambda t})^l z^l.$$

Thus:

$$G(z,t) = \sum_{\ell=N_0}^{\infty}\binom{\ell-1}{\ell-N_0} e^{-\lambda N_0 t}(1-e^{-\lambda t})^{\ell-N_0} z^{\ell}. \tag{4A.8}$$

The probability that the number of final goods at t is N, $Q(N,t)$ is the coefficient of z^N of this generating function (4A.8), and is given by:

$$Q(N,t) = \binom{N-1}{N-N_0} e^{-\lambda N_0 t}(1-e^{-\lambda t})^{N-N_0}.$$

Without loss of generality, we can take N_0 as 1. Then $Q(N, t)$ becomes:

$$Q(N,t) = e^{-\lambda t}(1-e^{-\lambda t})^{N-1}.$$

This is equation (4.8) in the main text, and is called the negative binomial distribution.

NOTES

* Reprinted with permission from *Journal of Economic Behavior and Organization*, **48**, (2002), copyright 2002 by Elsevier Science B.V. We thank seminar participants at Tohoku, Osaka, Keio, Siena and Tokyo universities. We want to give special thanks to Professors Kazuya Kamiya, Yoshiyasu Ono, Etsuro Shioji and Robert M. Solow who read the chapter very carefully and provided us with very useful comments and suggestions. We are also grateful to two anonymous referees for their extremely detailed comments which led us to a major revision of the chapter. All remaining errors are, of course, ours. The second author gratefully acknowledges financial support of the Ministry of Education, Grant in Aid for Scientific Research (No. 10430002) and the Economic Planning Agency, International Forum of Collaboration Projects.

1. This point trivially applies to the case where the quality improvement occurs in the production of intermediate goods. It also applies to the case where the 'creative destruction' or the quality improvement occurs in the production of final consumables. See, for example, equation (21) of Caballero and Jaffe (1993). If the number of commodities remains constant ($N=0$), the growth rate of the economy is $\hat{\eta}$, the growth rate of labor productivity in the production of the existing commodities.
2. Kuznets (1953), for example, argues: 'In the industrialized countries of the world, the cumulative effect of technical progress in a number of important industries has brought about a situation where further progress of similar scope cannot be reasonably expected. The industries that have matured technologically account for a progressively increasing ratio of the total production of the economy. Their maturity does imply that economic effects of further improvements will necessarily be more limited than in the past'. Based on the American experiences, McLaughlin and Watkins (1939) share this kind of pessimism.
3. Thus, the present analysis following Arrow (1962) and Stokey (1988) abstracts from profit maximization in R&D.
4. Yoshikawa (1995) explains the high growth of the Japanese economy during the 1950s and 1960s, emphasizing diffusion of consumer durables.
5. Nelson and Pack (1999) discuss only the general importance of productive assimilation. From the perspective of the present chapter, Asian NICs could grow rapidly because they succeeded in productive assimilation in those sectors where growth of demand was very high.
6. See Chapter 2 of Yoshikawa (1995).

REFERENCES

Abramovitz, Moses and I.A. Stegun (1968), *Handbook of Mathematical Functions*, New York: Dover.

Aghion, P. and P. Howitt (1992), 'A model of growth through creative destruction', Econometrica, **60**, 323–51.

Aoki, Masanao (1996), *New Approaches to Macroeconomic Modeling*, Cambridge: Cambridge University Press.

Aoki, Masanao (1997), 'Shares in emergent markets: dynamics and statistical properties of equilibrium classification of agents in evolutionary models', in T. Katayama and S. Sugimoto (eds), *Statistical Methods in Control and Signal Processing*, New York: Marcel Dekker.

Arrow, K. (1962), 'The economic implications of learning by doing', *Review of Economic Studies*, **29**, 155–73.
Caballero, R. and A. Jaffe (1993), 'How high are the giant's shoulders: an empirical assessment of knowledge spillovers and creative destruction in a model of economic growth', in O. Blanchard and S. Fischer (eds), *NBER Macroeconomics Annual 1993*, Cambridge, MA: MIT Press, pp. 15–74.
Diamond, P. (1982), 'Aggregate demand management in search equilibrium', *Journal of Political Economy*, **90**, 881–94.
Dutt, A.K. (1986), 'Growth, distribution and technological change', *Metroeconomica*.
Gates, Bill (1999), *Business@The Speed of Thought*, New York: Warner Books.
Gordon, Robert (1999), 'Monetary policy in the age of information technology: computers and the Solow paradox', IMES Discussion Paper, no. 99-E-12, Bank of Japan, Tokyo.
Grossman, Gene M. and E. Helpman (1991), *Innovation and Growth in the Global Economy*, Cambridge, MA: MIT Press.
Hoppe, F.M. (1984), 'Polya-like urns and the Ewens sampling formula', *Journal of Mathematical Biology*, **20**, 91–4.
Horowitz, A. and E. Lai (1996), 'Patent length and the rate of innovation', *International Economic Review*, **37**, 785–801.
Jones, Charles (1995), 'R&D-based models of economic growth', *Journal of Political Economy*, **103**, 759–84.
Jones, Charles and John Williams (1998), 'Measuring the social return to R&D', *Quarterly Journal of Economics*, **113**, 1119–35.
Kaldor, N. (1957), 'A model of economic growth', *Economic Journal*, **68**, 591–624.
Keynes, J.M. (1936), *The General Theory of Employment, Interest, and Money*, London: Macmillan.
Kuznets, Simon (1953), *Economic Change*, New York: Norton.
Marglin, S. (1984), *Growth, Distribution, and Prices*, Cambridge, MA: Harvard University Press.
McLaughlin, G. and R. Watkins (1939), 'The problem of industrial growth in a mature economy', *American Economic Review*.
Montroll, E.W. (1978), 'Social dynamics and the quantifying of social forces', *Proceedings of National Academy of Sciences* **75** (10).
Murphy. K., A. Shleifer and R.Vishny (1989), 'Industrialization and the big push', *Journal of Political Economy*, **97**, 1003–26.
Nelson, R. and H. Pack (1999), 'The Asian miracle and modern growth theory', *Economic Journal*, **109**, 416–36.
Robinson, J. (1962), *Essays in the Theory of Economic Growth*, London: Macmillan.
Romer, Paul (1990), 'Endogenous technological change', *Journal of Political Economy*, **98**, 71–1020.
Rostow, W.W. (1960), *The Process of Economic Growth*, Oxford: Oxford University Press.
Rostow, W.W. (1978), *The World Economy: History and Prospect*, Austin, TX: University of Texas Press.
Schmookler, Joseph (1966), *Invention and Economic Growth*, Cambridge, MA: Harvard University Press.
Schumpeter, Joseph (1934), *Theory of Economic Development*, Cambridge, MA: Harvard University Press.
Segerstrom, Paul (1998), 'Endogenous growth without scale effects', *American Economic Review*, **88**, 1290–310.

Smith, K. (1969), 'The effect of uncertainly on monopoly price, capital stock and utilization of capital', *Journal of Economic Theory*, **1**, 48–59.
Solow, Robert M. (1956), 'A contribution to the theory of economic growth', *Quarterly Journal of Economics*, **70**, 65–94.
Solow, Robert M. (1997), 'Is there a core of usable macroeconomics we should all believe in?', *American Economic Review*, **87**, 230–32.
Stokey, N. (1988), 'Learning by doing and the introduction of new goods', *Journal of Political Economy*, **96**, 701–17.
Uzawa, Hirofumi (1969), 'Time preference and the penrose effect in a two-class model of economic growth', *Journal of Political Economy*, Part II, **77**, 628–52.
Wright, Gavin (1997), 'Towards a more historical approach to technological change', *Economic Journal*, **107**, 1560–66.
Yoshikawa, Hiroshi (1995), *Macroeconomics and the Japanese Economy*, Oxford: Oxford University Press.
Young, A. (1995), 'The tyranny of numbers: confronting the statistical realities of the East Asian growth experience', *Quarterly Journal of Economics*, **110**, 641–80.
Young, A. (1998), 'Growth without scale effects', *Journal of Political Economy*, **106**, 41–63.

PART II

Demographic Transition and Pension Systems

5. Distributional impact of social security reform

Barry Bosworth, Gary Burtless and Claudia Sahm*

1. INTRODUCTION

The United States, along with most other industrial countries, is debating how to modify public retirement programs in response to population ageing. The debate has identified three broad approaches to reform: increasing contribution rates; reducing benefits; and pre-funding a larger fraction of future obligations. Opinions about these approaches differ because the options have differing distributional impacts, both on high- and low-wage workers within a cohort and across age cohorts. The debate over the first two policy options is dominated by distributional concerns. Boosting contribution rates will favor workers who are already retired or near retirement; reducing benefits hurts people who are retired or near retirement. The policy choice between the two options is viewed from the perspective of a zero-sum conflict in which the benefits or taxes of one generation or group of workers must be sacrificed in the interest of maintaining the incomes of another. The total amount of future resources available for consumption is assumed fixed, and the debate is over how to divide them between workers and retirees and between high- and low-wage workers.

From an economic perspective, the option of advance funding introduces a different kind of choice. The pool of resources for future consumption cannot be assumed to remain constant. It can rise or fall depending on today's choice of an advance funding policy. Current workers can fund a greater part of the cost of their own pensions by increasing their contributions to a retirement plan, and this is true whether the plan is publicly or privately managed. If workers' contributions are saved and used to finance the accumulation of additional capital, the result will be an expansion of the resources available to divide among future workers and retirees.

In earlier research we constructed a neoclassical growth model to assess the macroeconomic benefits of a strategy of funding a larger fraction of the future liabilities.[1] Our analysis suggested that a relatively small increase in funding, if it resulted in increased saving and investment equivalent to about 1 percent of GNP, could boost the growth of the economy enough to offset all of the added costs to future workers of supporting a larger retired population. That is, while future workers' pension contributions would still need to rise to cover larger benefit payments, the increase in future wages would be large enough to increase future workers' after-tax income.

A policy of increased saving has other macroeconomic consequences relevant to the debate over pension reform, however. In particular, a policy of investing the funds domestically will drive down the rate of return to capital, thus reducing the investment earnings of any retirement fund. The magnitude of the decline is potentially large because any future increase in saving will occur against a backdrop of much slower growth in the American labor force, which in turn will slow the domestic demand for capital. Alternatively, the extra saving could be invested in the larger global economy where the demand for capital is likely to remain high. If all the increase in saving were invested abroad, however, the impact of higher saving on future US wages and pension contributions would be small.

The issue of the dynamic effects of pension reform on the economy and the pension system takes on particular relevance when we consider the distributional consequences of reform. Reform proposals that aim to boost national saving can easily induce increases in future wages that dwarf the direct consequences of changes in contribution rates or the benefit formula. In a dynamic context, some reform proposals may also have implications for the financial condition of the social security fund, because they induce changes in future benefit payments and investment earnings as well as tax revenues. The induced effect on the rate of return can have particularly large implications for defined-contribution pension plans.

In this chapter we examine the distributional consequences of three alternative approaches to pension reform within a model that takes account of the dynamic effects of the reforms on aggregate output, wages and interest rates. This framework allows us to measure the feedback effects of advance funding on workers' earnings and pension benefits. These dynamic effects are sufficiently large so that they fundamentally change the conclusions we would draw based on a static analysis. Our analysis is performed by combining a small neoclassical growth model with a microsimulation model. This combination permits us to measure the impact of alternative reforms on a representative set of individual workers. The microeconomic model is based on a small number of earnings patterns that reflect the diverse career wage profiles of recent American workers. The profiles allow us to calculate individual pension benefits, lifetime net incomes and internal rates of return, and thus to examine the distributional impact of reform.

We focus on three alternative reforms that would restore social security to solvency: (i) future tax increases that maintain the current Old-Age, Survivors and Disability Insurance (OASDI) system on a pay-as-you-go basis (that is, each year's income roughly covers the year's expenses while maintaining a reserve equal to one year's expenditures); (ii) tax increases that maintain the system in actuarial balance over a 75-year horizon (a reform that requires a sizeable *immediate* tax increase); and (iii) benefit reductions that maintain solvency on a pay-as-you-go basis combined with a new mandatory system of individual pension accounts that partially offset the cuts in OASDI benefits. We have designed the second and third reforms so that both require workers to contribute an additional 2.6 percentage points of taxable wages to the retirement system starting in 2000. Under the second reform, the extra contributions flow into the OASDI trust funds. Under the third reform, they flow into private retirement accounts in which 70 percent of accumulations are held as equities and 30 percent are held as government bonds.

We evaluate these reforms first within a static framework in which the changes in pension fund contributions and benefits have no effect on saving, capital formation or future pre-tax wages. The microeconomic portion of the model is used to compute future contributions, benefits and internal rates of return under each of the reform rules. The impacts of the reforms are shown for workers with differing earnings profiles within a given age cohort. They are also calculated for 81 successive age cohorts retiring between 1995 and 2075.

In Section 3 we report the results of a full dynamic simulation in which pension reforms lead to increases in national saving, domestic investment and the level of future wages and benefits. Induced increases in future wages come at the expense of lower rates of return on capital. To assess the impact of these changes on individual workers, we show the effect of the reforms on workers' lifetime income, measured as the sum of after-tax earnings and pension benefits. We argue that standard measures for assessing the individual gains from pension reform, such as changes in the internal rate of return on contributions or pension replacement rates, do not adequately reflect changes in worker well-being. Some of the dynamic feedback effects of reform are missed by these measures. The appropriate benchmark for gauging workers' gains or losses under a reform is the change in their lifetime income. Reforms that preserve or even boost workers' returns on their pension contributions may be ones that fail to improve their lifetime income. Reforms that produce a low apparent return on workers' contributions by driving down the return on capital may produce sizeable gains in their lifetime income by increasing before-tax wages.

2. STATIC ANALYSIS

We begin by describing the sample of workers used to infer the distributional effects of reform. We then describe the macroeconomic model and trust fund accounting model used to calculate the size of tax increases and benefit cuts that are needed to keep the social security system solvent. Crucial statistics about the future economy and population are derived from the intermediate economic and demographic assumptions of the *1996 OASDI Trustees' Report* (Board of Trustees, OASDI 1996). Others are generated by a standard neoclassical growth model. The microsimulation model consists of a set of stylized age–earnings profiles and algorithms for calculating the earnings, taxes and retirement benefits of individual workers under each of the policy alternatives. The individual profiles are directly linked to predictions from the macroeconomic model. Each worker's annual earnings at a given age can be expressed as a percentage of the economy-wide average wage. In the baseline simulation, we assume that future economy-wide wages follow the exact path predicted in the *1996 OASDI Trustees' Report*.

Age–Earnings Profiles

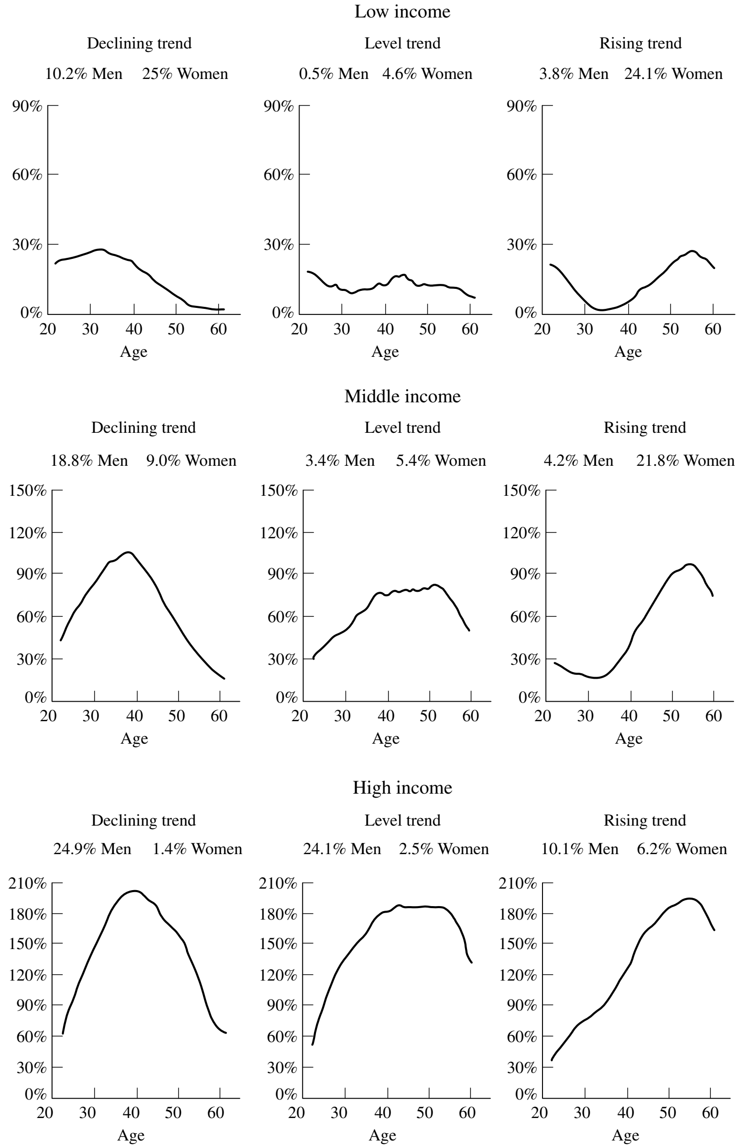
The distributional impact of reform can only be determined by measuring the effects of the policy change on representative individual workers. We calculate reform impacts on nine representative profiles of American workers. The specific earnings profile of each worker was estimated with information from the US Census Bureau's 1990–93 Survey of Income and Program Participation (SIPP) matched to social security earnings records (SSER).[2] The SSER records contain information on social-security-covered earnings by calendar year for the period from 1951 to 1996. The profiles of relative earnings are based on the observed earnings of all social-security-covered workers born between 1931 and 1940. The calculations were performed on a sample that included all workers in the sample with at least one year of covered earnings. We measure earnings at each age relative to the economy-wide average wage in that year as reported by the Social Security Administration (SSA).

Individuals are classified into nine categories based on the average *level* of their relative earnings over their career (low, middle and high) and the *trend* in their career earnings profile (declining, level and rising). For purposes of calculating workers' retirement benefits under the current social security formula, it is enough to know the level of their career wages (specifically, the highest 35 years of indexed earnings). On the other hand, the trend or time path of earnings has a major impact on benefits under a defined-contribution (DC) pension plan. Contributions into a DC account in the early years of a worker's career earn investment returns over a longer period, providing a larger pension per dollar contributed than contributions made late in the career.

For each of the approximately 25 000 workers in our original sample, we measured career earnings using observed earnings in the 40 years between ages 22 and 61. To estimate the average level and trend in career earnings, we focused on the last 30 years of the career.[3] We further divided this 30-year period into three 10-year subperiods. The average 'level' of a worker's earnings is the simple average of relative earnings over the 30 years. The 'trend' is defined as the direction of change in 10-year average earnings between the first and third decades of the three-decade period.[4]

The nine estimated age–earnings profiles are displayed in Figure 5.1 (with y-axis scales adjusted to reveal greater detail). Each chart in the figure shows a single age–earnings profile. At a given age, the chart shows the worker's annual earnings measured as a percentage of US economy-wide earnings at that age. The figure shows a remarkable diversity of earnings patterns. As many workers have a declining trend in career earnings as have a rising profile. The percentages of men and women in each category are shown at the top of the figures. As expected the distribution of women's earnings is lower than that for men, but women are more likely to have a rising trend over their career. In general, the shape of the age–earnings profile of men and women *within* a category is very similar. The important difference between the two sexes is in the distribution of workers across the nine categories.

These representations of low, middle and high earners differ significantly from the traditional stylized earnings patterns used by the SSA.[5] Our tabulations suggest that workers in the bottom third of the lifetime earnings distribution have an average wage that is less than half the low-wage value assumed by the SSA. Our estimates for the middle- and high-wage groups suggest that lifetime earnings are about two-thirds of the amounts assumed by the SSA for corresponding groups of workers (see Bosworth et al. 2000). One reason for these differences is that our estimated earnings profiles include years in which individual workers may have no earnings. This seems appropriate for the purpose of computing retirement benefits. Over the full 40-year career the rate of nonemployment is about 50 percent for workers in the low-wage group. That is, the low-wage workers have positive earnings amounts in only about half of the 40 years between ages 22 and 61. The nonemployment rate is less than 10 percent for the high-wage workers. If we ignored years with zero earnings in our tabulations, it would boost the average level of the earnings profiles, especially of low-wage workers, but it would have relatively little influence on the shape of the individual profiles.



Note: Earnings measured as a percent of economy-wide average earnings.

Source: Authors' tabulations of matched SIPP-SSER files (1990–93 SIPP panels).

Figure 5.1 Age–earnings profiles of nine clases of workers, 1931–1940 birth cohorts

Our tabulations include only those workers who survive to age 62. In the distributional analysis we assume workers apply for a pension when they reach 62. After age 62, we use the intermediate mortality assumptions of the 1996 Social Security Trustees' Report to calculate composite future survival rates. We also assume, unrealistically, that future life tables will be identical for male and female workers as well as for workers in all nine earnings profiles.[6] While we ignore the effects of gender and average wages on mortality rates, our estimates incorporate the predicted effects of future improvements in longevity for successive cohorts.

Using the age–earnings profiles displayed in Figure 5.1 and our assumptions about future survival rates, it is straightforward to calculate the value of future pensions under the existing social security benefit formula and a variety of revised formulas. We can thus calculate the effect of a reform on the after-tax earnings and benefits of representative workers who attain 62 and retire in successive years from 1995 to 2075.

Alternative Solvency Rules

The present social security system is insolvent. The accumulated Trust Fund reserve plus anticipated tax contributions and interest earnings are not large enough to pay for the accumulated and future liabilities of the system. Under the intermediate economic and demographic assumptions of the 1996 Trustees' Report, which we use here, the Trust Fund was predicted to be exhausted in 2029.[7] In one sense, this means our baseline forecasts of future OASDI revenues and outlays are grossly unrealistic, because they do not address the funding shortfall that will occur after 2029. To underscore the inconsistency of the projections, by the end of the forecast period in 2075 the cumulative OASDI shortfall would be approximately 100 percent of GDP in that year. Such a shortfall is not likely to be tolerated by voters or policymakers. All three reform scenarios described below require that the long-term insolvency in the social security system be addressed, either through tax increases or benefit cuts.

Policy option 1 requires that solvency be achieved entirely through tax increases, with the yearly tax rate calibrated so as to maintain solvency under a modified pay-as-you-go rule. In each year of the forecast period (2000 to 2075), our rule requires that anticipated tax revenues in the year plus reserves held in the Trust Fund be sufficient to pay for at least two years' anticipated outlays. If revenues, outlays and Trust Fund reserves do not pass this test, we increase the payroll tax by the minimum amount needed so that the test is passed. In years when the Trust Fund contains reserves equal to 100 percent of anticipated outlays and when the OASDI cost rate is constant or increasing, the required tax is almost identical to a pure pay-as-you-go tax rate (that is, the tax rate is just sufficient to pay for current benefit payments and administrative costs). In years when the Trust Fund contains significantly more than 100 percent of annual outlays, the tax rate may fall short of the pay-go tax rate. To eliminate the possibility of tax-rate fluctuations, we do not allow the payroll tax rate to fall in any future year.[8] The size and timing of the necessary tax increases under option 1 are shown in Figure 5.2. The social security program first fails our modified pay-go test in 2024. Between 2024 and 2034 the OASDI tax rate must increase from 12.4 percentage points to 16.6 percentage points, implying that the OASDI tax burden will increase by one-third in 10 years. The

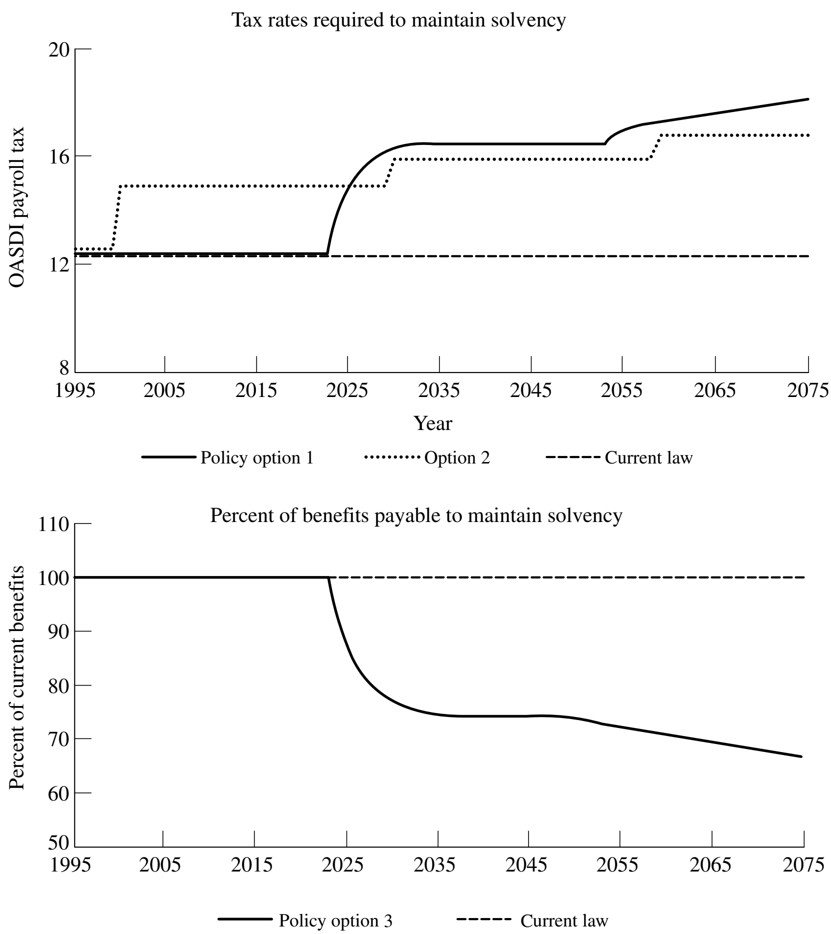


Figure 5.2 Solvency adjustments in OASDI tax rates and benefits

rise in payroll tax rates is ultimately due to the increase in social security benefit outlays as the baby-boom generation retires. The abruptness of the increase can be traced to the accumulation of Trust Fund reserves between 1985 and 2015. The reserve accumulation during these three decades is so large that the payroll tax can remain below the OASDI cost rate for a number of years without triggering a payroll tax hike under our pay-go rule. Once the tax hikes begin, however, large increases are needed to keep the program solvent. A key point is that much of the cost of restoring social security solvency is deferred until after most workers who are active in 2000 have already retired. By 2075 the required pay-go tax rate must rise to 18.3 percent, almost 6 percentage points higher than the rate scheduled under current law.

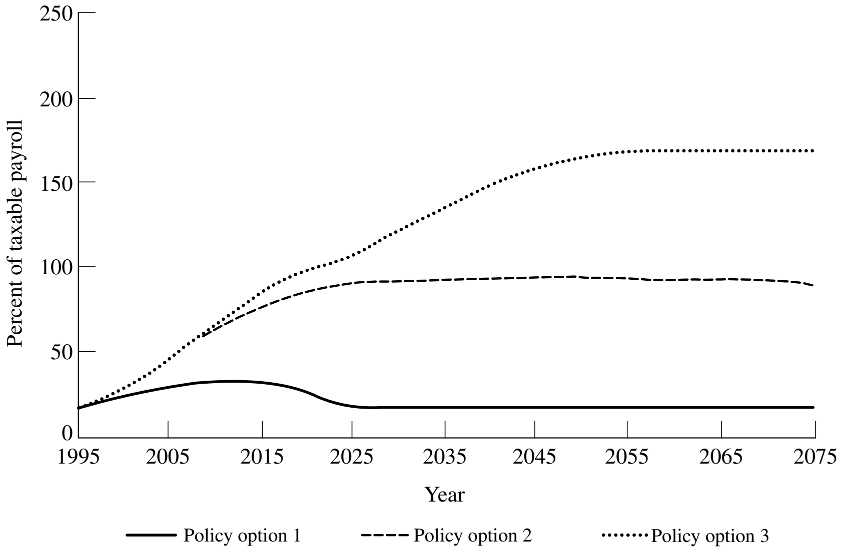
Policy option 2, in contrast, requires workers active in 2000 to contribute more toward their future benefits through advance funding of OASDI obligations. As under option 1, we assume that the current benefit formula will be maintained. Under option 2, however, the payroll tax is periodically adjusted to maintain close actuarial balance over a 75-year horizon, the longest planning horizon generally used in the Trustees' Reports. A 75-year solvency rule essentially requires that the 'summarized income rate' of OASDI be at least 95 percent of the 'summarized cost rate' of the program, where the income and cost rates are measured at the beginning of the 75-year period. OASDI income consists of OASDI income over the planning period plus the initial reserves in the Trust Fund. The summarized income rate is the discounted OASDI income divided by the discounted taxable payroll. OASDI costs consist of 76 years of benefit payments and administrative costs discounted to the present. (Seventy-six rather than 75 years of outlays are summed because of the requirement that the Trust Fund at the end of the planning period contain enough funds to pay for one additional year of program costs.) The summarized cost rate is the discounted sum of costs divided by the discounted taxable payroll. All discounting is performed using the projected rate of return earned by the Trust Fund.

We consider 76 different planning periods, each beginning in successive years from 2000 to 2075. Whenever our calculations show that the summarized income rate in a planning period is less than 95 percent of the summarized cost rate, we raise the payroll tax rate by the amount needed to bring the summarized income rate up to 100 percent of the summarized cost rate in the period. The tax hike begins in the first year of the planning period and is assumed to remain in effect throughout the period. It may later turn out that the payroll tax rate must be increased again in a subsequent year, because the program fails the 75-year solvency test in a planning period that begins in a later year. Whenever the solvency test fails, the tax is increased.[9] The first tax hike required under this reform occurs in 2000, the first year of our policy simulation. OASDI is not in close actuarial balance in that year, and the payroll tax rate must be increased 2.6 percentage points to restore full actuarial balance.[10] The top panel in Figure 5.2 shows the tax increases needed to maintain solvency under the second solvency rule. Initially, the required tax increase produces a higher payroll tax rate than needed under the first policy option. Eventually the buildup of an interest-earning reserve reduces the required magnitude of tax increases after 2025. In that sense, the baby boom generation pays for a bigger portion of its own retirement, reducing the tax burden imposed on some future workers.

Policy option 3 scales back future social security pensions but offsets the benefit cut with the introduction of new individual retirement accounts. Social security benefits are periodically reduced to keep OASDI solvent on a pay-go basis using the same modified pay-go rule described for option 1. The important difference, of course, is that the current OASDI tax rate – 12.4 percent – is preserved instead of the current OASDI benefit formula. In the year 2000, all workers who are not within five years of retirement are assumed to begin contributing to the new accounts. We specified a contribution rate of 2.6 percent to correspond with the initial tax hike required under option 2, thus we are able to evaluate approximately equivalent funding schemes both within and outside the current social security system. The individual accounts are assumed to be invested 70 percent in equities and 30 percent in government bonds. In our baseline simulation we assume the bond rate is the same as assumed in the 1996 Trustees' Report. We estimated the baseline return on equity by assuming that the weighted average return on bonds and equities is equal to the after-tax return on physical capital. This assumption implies an equity risk premium of about 5 percent.

Whenever an OASDI benefit cut is needed to preserve solvency under our modified pay-go rule, the cut is imposed on all OASDI pensioners collecting benefits in that year.[11] The bottom panel of Figure 5.2 shows the time path of required benefit cuts under the third policy option. The timing of the cuts coincides with the tax increases needed under option 1, but the percentage changes differ. From 2024 to 2034 OASDI benefits must be reduced under option 3 by about 25 percent of the amount promised under the 1983 social security amendments, while under option 1 payroll tax rates must increase by almost a third. The percentage difference occurs because payroll taxes represent only part of the revenue available to pay social security benefits. (Income taxes on OASDI pensions are another source of funding.) Moreover, under our baseline assumptions benefit costs far exceed tax receipts. As a result, option 1, which alters only payroll taxes, requires larger percentage changes than the benefit cuts implied by option 3.

Figure 5.3 shows pension reserves measured as a percent of taxable



Note: Government dissaving is 0% of NNP after 1999. Additions to Trust Fund reserves do not add to national saving.

Figure 5.3 Pension reserves as a percent of taxable payroll in baseline

payroll under the three reforms. Under the first and third policy options, the reserves in the Social Security Trust Fund do not become particularly large. Under the first option, of course, there are no pension reserves outside the Trust Fund. Under the third option, workers also accumulate sizeable reserves in their individual accounts. The combined assets of OASDI and the individual accounts under option 3 grow rapidly up to about 2040, when workers who have contributed over their entire work life begin to retire. Thereafter the rate of accumulation slows and the pension fund stabilizes at 160 percent of taxable payroll. Under the second policy option, the reserve accumulated in the Social Security Trust Fund approaches 100 percent of taxable payroll, about midway between the reserve accumulations implied by the other two policies. By 2075 the level of reserves under option 2 is about half that of option 3. The difference arises primarily because the individual accounts are heavily invested in equities, which earn a higher return than the Treasury bonds held in the OASDI Trust Fund. Nonetheless, if the accumulated pension reserves add to overall national saving, it is clear that the increases to saving will be much larger under either of the advance-funding plans than under the pay-as-you-go option.

Distributional Consequences

Analysts have traditionally emphasized two measures to evaluate pension reforms, the replacement rate and the internal rate of return. The SSA typically calculates the replacement rate as the real pension at retirement divided by the real taxable wage just before retirement. Because most workers' earnings profiles do not exhibit smoothly trending real wages, this way of measuring replacement rates is not very meaningful for the nine earnings profiles displayed in Figure 5.1. A more appropriate measure of the replacement rate would use career-average wages in the calculation. More fundamentally, the replacement rate is not well suited to the evaluation of alternative reforms, since it does not accurately reflect the implications of tax changes. A tax increase has no effect on a replacement rate based on the before-tax wage. Perversely, a tax increase would actually improve the replacement rate if the ratio were measured using the after-tax wage.[12]

The internal rate of return is the discount rate that equilibrates the discounted value of real contributions and expected benefits over an individual's life. We provide estimates of the internal rate of return of the OASI program for the nine profiles of workers who retire at 25-year intervals in Appendix Table 5A.1. While the internal rate of return is useful in a static context for evaluating alternative reform options, it is seriously flawed when applied to situations where the reforms lead to changes in future pre-tax wages. The internal rate of return misses much of the benefit workers derive from wage gains induced by the reforms. An important rationale for advanced funding of pension liabilities is the desire to boost saving and potentially to increase future wages. It seems important to employ a measure that captures the effect of these changes on workers' net incomes.

For this reason we measure the impact of each reform option on workers' lifetime net incomes, taking account of changes in both after-tax earnings and expected pensions. In a static analysis, where before-tax wages are assumed to remain unchanged, this method of measuring reform impacts makes little difference. It is much more revealing for evaluating social security reform in a full dynamic simulation. Our calculation of lifetime earnings takes account of any change in payroll taxes and DC pension contributions in computing lifetime net earnings. In measuring lifetime pensions under option 3, we combine social security and DC pensions and assume that workers are required to convert to a fixed real annuity at age 62.[13] Lifetime income is then a simple cumulative sum, calculated at age 62, of past after-tax earnings and expected future pensions.

The impact of the three reforms on the lifetime incomes of retiring workers is displayed in Figure 5.4. The effect of a reform is measured relative to a

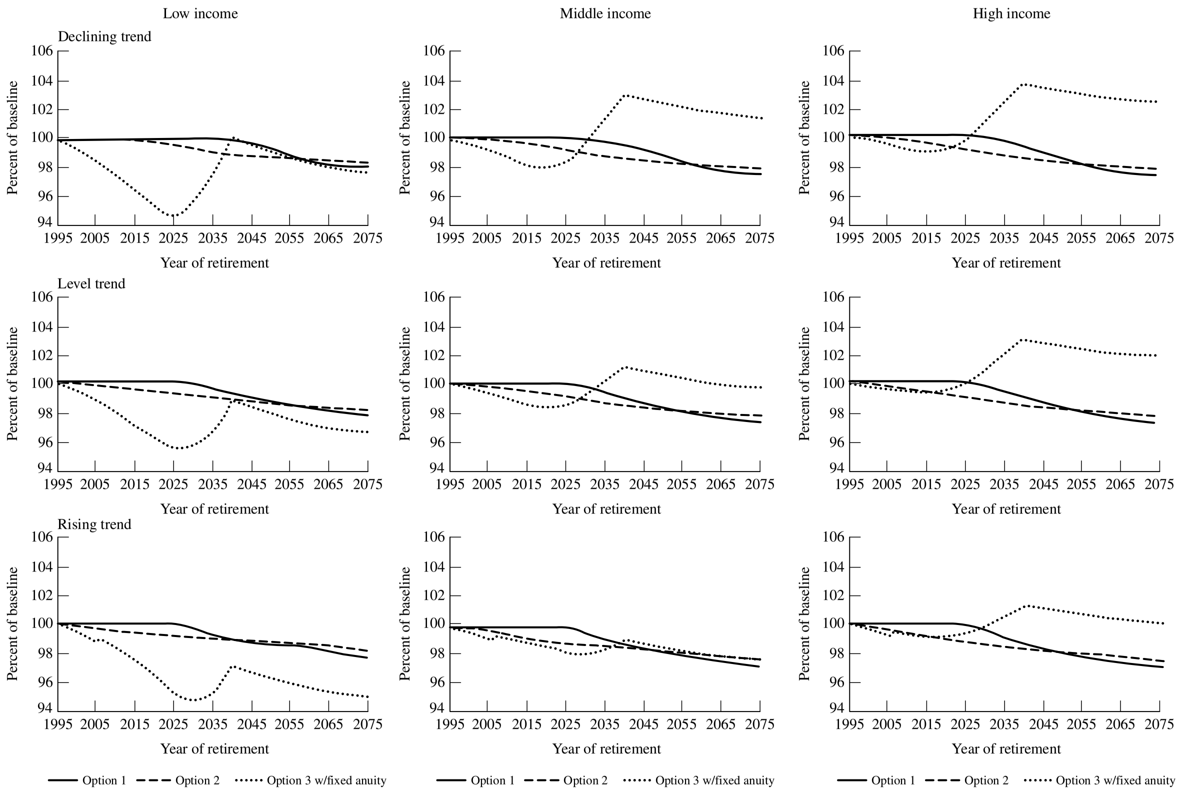


Figure 5.4 Total lifetime income as a percentage of the baseline values in a static simulation

worker's lifetime income in the baseline simulation. We have prepared one chart for each of the nine earnings profiles, and for each profile we have calculated the percentage change from the baseline lifetime income for workers retiring in successive years from 1995 to 2075. The change in income is reported for workers reaching age 62 in the year shown at the bottom of the chart. Additional detail is provided in Appendix Table 5A.2, where the lifetime earnings and pension components are shown separately.

The first reform maintains existing benefit levels with little advance funding, using a modified pay-go rule to determine required tax increases. The impact of the reform is negligible for workers who retire before 2024. Because the change imposed by the reform is limited to a proportional change in payroll taxes, workers in a given cohort all experience approximately the same percentage loss in lifetime income. Even more important, they are protected against further income loss after they retire. Because required tax increases grow ever larger after 2024, the lifetime income losses grow steadily bigger for successive cohorts. In effect, the reform protects the incomes of workers who retire before 2024 at the expense of workers who retire later. This implies substantial, inter-cohort transfers from the young to the old.

Under option 2 social security benefits are maintained, but tax increases occur much earlier than under the first plan. Consequently, early cohorts of retirees suffer larger lifetime income losses than under the first reform option. Later cohorts suffer smaller losses, however, because part of the burden of achieving solvency is borne by earlier cohorts. Thus option 2 reduces the intergenerational transfers. Neither of the two tax-increase plans alters the within-cohort distribution of lifetime income relative to the current system.

Option 3 produces the most interesting results. Because the benefit reductions in that plan apply to all beneficiaries, workers who retire in a given year are affected by cuts even if they retire before any benefit cut is imposed. Though the first benefit cut is delayed until 2024, the severity of the cuts after 2024 requires large retirement income sacrifices from workers who retire as early as 2015. The effects are particularly large for low-wage workers, because social security benefits represent a large percentage of their lifetime income. The loss of social security benefits is not offset by pensions from individual accounts, because workers retiring before about 2030 have not made contributions to the accounts during a long enough portion of their career. (We assume DC plan contributions begin in 2000.)

The size of income gain or loss also varies substantially across the nine earnings profiles. Because low-wage workers receive much better social security benefits (relative to contributions) than middle- and high-wage workers, it is hard for them to make up the losses of OASI benefits from pensions out of a DC account. It is also clear that the DC plan is particularly disadvantageous for workers who earn most of their lifetime earnings late in their career (see Appendix Table 5A.2). In contrast, the option is advantageous for high-wage workers who retire after a full career of contributions to the DC plan. For high-wage and most middle-wage workers who retire after about 2030, the introduction of an individual account plan boosts lifetime income. For most workers who retire before 2030 and for nearly all low-wage workers who retire before or after 2030, the individual account plan provides lower lifetime income than they could obtain under the two policies that maintain current social security benefits.

It would appear from this static analysis that options 1 and 2 have roughly similar implications for workers with differing average lifetime earnings, at least within the same birth cohort. Option 2 reduces modestly the degree of intergenerational transfers between the old and the young, but the magnitude of the change is dwarfed by that of option 3.

The interesting within-cohort distributional issues arise under option 3, which scales back current social security benefits and introduces a system of DC accounts. This substitution imposes large losses on low-wage groups, but it is eventually beneficial to workers with high earnings. Option 3 would also have differential impacts on women and men. Women's earnings are on average lower and tend to be more concentrated in the later years of their career than is the case among men. Note in Figure 5.1 that 52 percent of women have rising career earnings profiles compared with just 18 percent of men.[14] DC accounts are less helpful to workers who earn most of their wages late in their career, because contributions do not benefit from as many years of compounding. Furthermore, because the basic social security package is scaled back under option 3, women would not gain as much under the revamped system from spousal and survivor benefits as they do under the existing system.

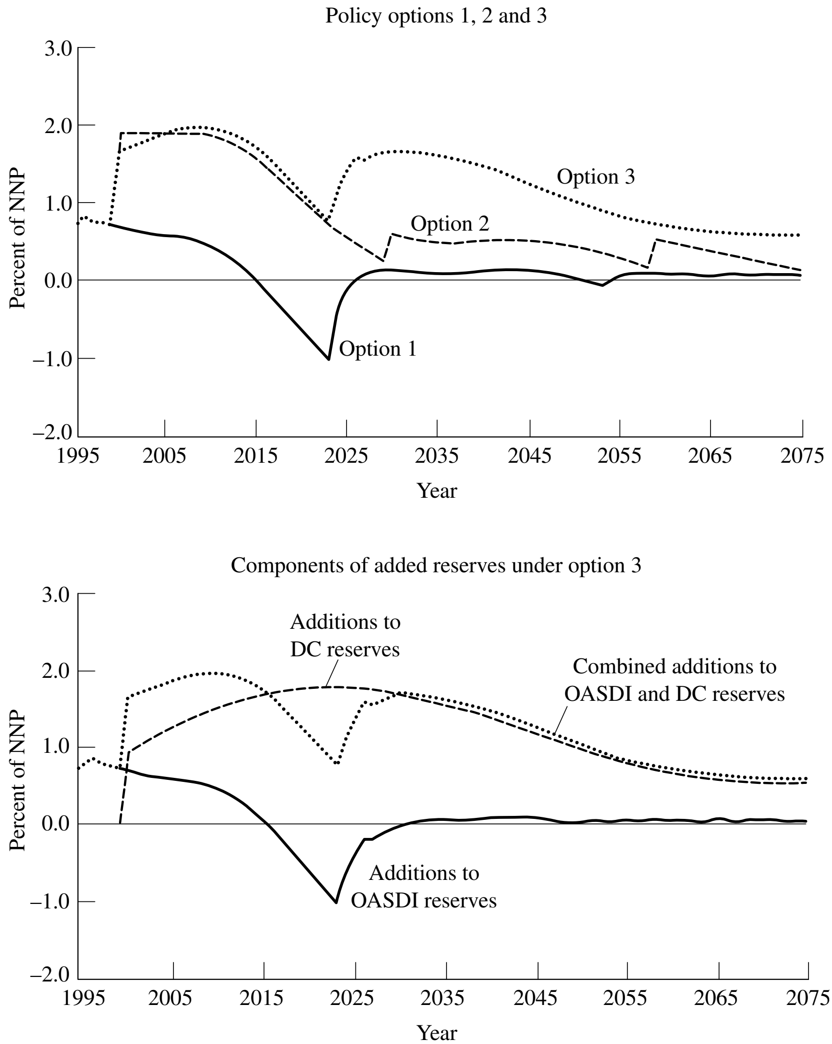
3. INCORPORATING THE EFFECTS OF REFORM ON NATIONAL SAVING, OUTPUT AND WAGES

The static analysis ignores the main rationale for reforms that fund a bigger portion of future pension obligations – their potential for increasing national saving and capital formation. A larger stock of capital can boost the pre-tax wages as well as pensions of future generations. There is no consensus, however, on the extent to which additional saving in pension accounts would be offset by lower saving in the public sector or in other private accounts.

It is sometimes claimed that the very existence of social security, by providing the promise of future retirement benefits, leads people to save less for their own retirement.[15] The appropriate question for our analysis is somewhat different, however. Would a move toward funding a larger percentage of *existing* benefit promises affect saving behavior? If the fund accumulation occurs within the public sector, some analysts are concerned that any surplus would be used to increase public spending or reduce non-social-security tax revenues. One question, then, is whether Congress could avoid the temptation of using a larger pension surplus for some other purpose.

Larger pension system saving could also be offset by lower saving in other private accounts. The impact on private saving probably depends upon the confidence that people attach to the promise of a social security pension. Individuals who did not previously anticipate receiving promised benefits might react to a stronger guarantee of OASDI solvency by reducing their own private saving. This scenario seems most plausible for workers who already have private retirement accounts of their own. In contrast, for workers with high confidence in future benefit promises, funding should have little impact on their own saving plans. Our model allows for a wide range of different assumptions about the public and private saving response to greater reserve accumulation. In the following analysis we adopt the extreme assumption that all net additions to pension fund reserves, after adjusting for inflation, add to net national saving. This provides a simple alternative to the equally extreme assumption embodied in the static analysis – that growing pension reserves have *no* impact on aggregate saving.

The potential implications of the three reforms for aggregate saving are suggested in Figure 5.5. These calculations show the annual additions to pension fund reserves, before the feedback effects on the capital stock and future economic activity have been taken into account. We define saving as the change in the real value of the funds' reserves.[16] Because the OASDI system begins with a surplus, option 1 would generate some additions to saving up until 2015. This is followed by a 10-year period of fund decumulation. Beginning in 2025, the increment to net saving essentially reaches zero under the pay-as-you-go rule. Option 2 generates a large increase in saving starting in 2000, when the payroll tax is increased by 2.6 percentage points. The contribution to saving ultimately shrinks as the Trust Fund stabilizes at a constant share of taxable wages. Because the mandatory contribution rate into new individual accounts under option 3 is set equal to the initial tax increase under option 2, the two proposals have similar implications for aggregate saving out to 2020. Once the OASDI trust fund stabilizes, however, option 3 begins to generate larger additions to net saving, primarily because reserves in the individual accounts are invested in assets with a higher yield (more equities and fewer government bonds).



Note: Government dissaving is 0% of NNP after 1999. Additions to Trust Fund reserves do not add to national saving.

Source: Authors' tabulations.

Figure 5.5 Additions to net savings from pensions in a static simulation

The macroeconomic effects of the additions to saving are shown in Table 5.1.[17] We assume that the government budget, excluding social security, is in balance and the private saving rate follows the same path as the baseline projections. In addition, the increments to saving are invested in the domestic economy.[18] As expected, the first policy option generates little or no change in future output or wages. The feedback effects of the plan are trivial. In contrast, both options 2 and 3 generate large increases in the future capital stock, GDP and wages. In both cases, the economy-wide wage rate is about 12 percent above the baseline projection after 25 years and is 15 percent higher after 75 years. The increase in saving implies a lower level of aggregate consumption in the first decade after 2000, but the additions to the capital stock and real wages allow real consumption to rise 4 percent above the baseline by 2025 and 6 percent above the baseline path by 2075. The increases in payroll tax receipts that result from pre-tax wage growth delay both the tax increases required under option 2 and the benefit cuts needed in option 3. However, because pre-tax wage gains eventually

Table 5.1 Effects on the economy of alternative social security reforms dynamic simulation[a]

Year	Capital services	GNP	NNP	Consumption	Wage rate	Return on capital	Net addition to saving[b]
Option 1: Pay-go tax increase							
2000	0.7	0.0	0.0	−0.9	0.0	−0.3	0.7
2025	2.2	0.7	0.7	1.1	0.7	−1.2	−0.4
2050	5.1	1.1	0.7	0.5	1.4	−4.7	0.1
2075	6.2	1.4	0.9	0.7	1.7	−5.8	0.1
Option 2: Pre-funding tax increase							
2000	1.9	0.0	0.0	−2.6	0.0	−0.8	2.0
2025	49.2	9.0	5.4	4.0	11.6	−32.6	1.0
2050	60.6	11.2	6.5	5.3	14.0	−40.2	0.7
2075	60.7	11.7	6.6	6.1	14.1	−41.9	0.2
Option 3: Pay-go benefit cut + 2.6% DC pension							
2000	1.7	0.0	0.0	−2.3	0.0	−0.7	1.8
2025	48.2	8.9	5.4	4.3	11.5	−32.1	0.7
2050	66.6	12.1	6.9	5.6	15.2	−43.0	0.7
2075	68.9	12.9	7.1	6.1	15.7	−45.9	0.4

Notes:
[a] Percent change in comparison with baseline unless otherwise noted.
[b] Net addition to saving is measured as a percent of contemporaneous net national product (NNP) in the baseline.

result in real social security benefit increases, the net impact of the larger capital stock on the financial position of the OASDI fund is modest.

The increased capital accumulation drives down the rate of return on capital. In options 2 and 3 the real rate of return is cut by nearly a third after 25 years and by almost half after 75 years. This is of particular significance for retirement incomes under option 3. The high assumed rate of return on assets held in the individual accounts was an important source of much of the extra saving accumulation under that plan compared with the second reform plan. In our model, a decline in the return to physical capital is divided between the bond and equity yields in proportion to the relative yields of the two kinds of asset. For each one-percentage-point decline in the return to physical capital, the bond rate falls by 0.3 points and the equity yield by 0.9 points. The decline in the rate of return thus falls disproportionately on the equity yield, sharply reducing the buildup within the individual accounts. As a result, options 2 and 3 ultimately generate roughly equivalent increases in national saving, and thus produce similar improvements in future wages.[19]

Figure 5.6 shows the distributional consequences of the reforms after macroeconomic feedback effects are taken into account. First, note the striking contrast with the results of the static analysis (presented in Figure 5.4). In the dynamic analysis, the two reforms that boost pension accumulations and national saving generate large additions to future lifetime incomes for all workers. The gains grow over time, reversing the pattern of intergenerational transfers implied by the static analysis. This finding reflects the fact that most of the gains from a policy of increased saving are obtained in the form of higher pre-tax wages, an impact that is missed in the static analysis which reflects only the impacts of tax-rate changes and pension cuts.

Note that the pattern of net gains to workers in the middle- and high-wage groups is very similar under both options 2 and 3. This also contrasts with the results of the static analysis, which imply that the individual account plan eventually provides higher retirement incomes to high- and middle-wage workers. In the dynamic simulation the large reduction in the return to capital reduces pension incomes flowing from the individual accounts under option 3, partially offsetting the improvements in workers' wages. (Additional detail about the impact of reforms on earnings and pension income is presented in Appendix Table 5A.3.) The reduction in the rate of return has a much smaller impact on retirement incomes under option 2, because social security benefits are linked to economy-wide real wages, rather than to the real rate of return.

The largest differences between options 2 and 3 are reflected in the results for low-wage workers. Because these workers earn a generous return under

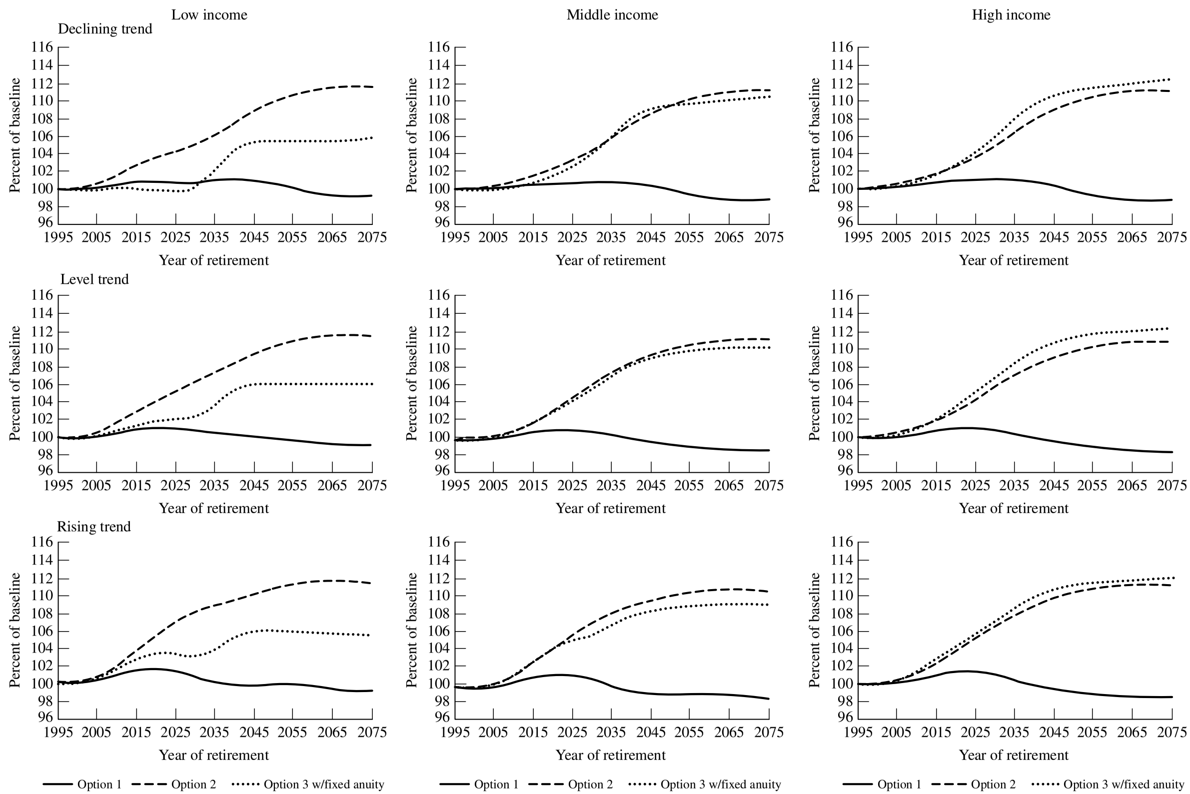


Figure 5.6 Total lifetime income as a percentage of the baseline values in a dynamic simulation

the social security benefit formula, their retirement income is higher under the wage-linked pension system of option 2 than under the DC pension accounts offered under option 3. Under both funding options, however, low-wage workers benefit noticeably in the form of higher wages.[20] Thus, unlike the conclusions of the static analysis, these two options generate gains to workers at all wage levels. Even though option 3 ultimately raises the lifetime income of all workers, it does it through a mechanism that forces all low-wage workers and some middle-income workers to accept large cuts in the proportion of their income that is received in retirement. Option 2 is more successful in preserving the retirement income of low-wage workers.

The results for option 1 are essentially the same as in the static case. Except for an effort to set aside the OASDI surpluses immediately after 2000, the plan has no significant effect on national saving. Thus, the option continues to impose costs on future generations that only seem larger in contrast to the gains these generations enjoy under options 2 and 3.

If pension reform yields an increase in national saving that boosts wages and lowers the return on capital, the internal rate of return that workers obtain on their pension contributions no longer provides a reliable guide to the gains they derive from reform. Table 5.2 illustrates this point for selected workers who retire in 2045. Each row in the table refers to a worker with a particular profile of career earnings. The bottom row refers to a worker with the composite earnings profile, that is, the weighted average profile of all workers in our sample. The first three columns show real internal rates of return on workers' pension contributions under the three reform options

Table 5.2 Internal rates of return and changes in lifetime income under alternative reform options for workers retiring in 2045, by career earnings profile

	Real internal rate of return on pension contributions (%)			Change in lifetime income (% of net income in baseline)		
Career earnings profile	Option 1	Option 2	Option 3	Option 1	Option 2	Option 3
Low wage/declining trend	4.6	4.3	3.8	0.9	9.1	5.3
Average wage/rising trend	3.5	3.4	3.1	−0.7	10.0	8.9
High wage/level trend	1.8	1.7	1.8	−0.2	9.3	10.8
Composite wage profile	2.8	2.6	2.5	−0.1	9.4	9.5

Source: Authors' tabulations (see text).

we have been considering. The last three columns show the change in workers' lifetime incomes, taking account of both net wage and pension changes, measured as a percent of each worker's lifetime income in the baseline simulation.

Judging by the return that workers obtain on their contributions to the pension system, option 1 clearly dominates the other two options. For each earnings profile, the real return under option 1 is at least equal to – and usually greater than – real returns under either of the other plans. Yet option 1 offers meager income gains in comparison with reforms that increase national saving. A worker with the composite wage profile receives 9.5 percent more lifetime income under the two plans that require more advance funding. Workers who retire in 2045 can anticipate making substantially larger lifetime contributions if the pension system moves toward more advance funding. If the extra contributions add to national saving and domestic investment, the higher flow of investment will increase the size of the capital stock, increasing productivity and wages, which further increases workers' required contributions to the pension system. A larger capital stock also depresses the return on capital. Because pensions are partly financed out of the interest earnings of the pension fund, the lower return on capital also pares workers' returns on their pension contributions. A plan that is beneficial for workers' incomes may be harmful for the rate of return, but it is nonetheless beneficial for workers' welfare.

4. CONCLUSION

The dynamic analysis illustrates the importance of incorporating the feedback effects of pension reform on the aggregate economy. If pension reform does not lead to changes in national saving, the choice of reform options is dominated by the distributional trade-offs – between the young and the old and between workers with high and low career earnings. There are no significant income gains to divide among current and future workers. If instead reform generates higher national saving, the dynamic effects of reform could include an increase in wages and a decline in the return to capital. Both of those changes have important implications for the comparison of social security reform options. The increase in economy-wide wages contributes to the well-being of future workers regardless of their income level. This is an important benefit of policy options that generate additional funding of future pension obligations.

The same process that yields improvements in economy-wide wages reduces the return on capital. The decline in the rate of return has a disproportionate impact on the gains of shifting from a defined-benefit system such as social security to a defined-contribution system of individual accounts. In the dynamic analysis, options 2 and 3 become very similar in terms of their implications for the well-being of the average worker. Both the dynamic and static simulations imply that low-wage workers are better off when social security benefits are maintained under option 2 than they would be if the present system were partially replaced with a DC plan. But in contrast to the static analysis, the dynamic results imply that middle- and high-wage workers are eventually about equally well off under both funding options. Our findings also highlight a weakness of any policy analysis that focuses solely on the rate of return workers obtain on their pension contributions. By emphasizing retirement income, such an analysis misses the impact of a policy on future pre-tax wage income. Yet for most workers the impact of a policy change on pre-tax wages is far more important in determining lifetime income than is the rate of return they obtain on pension contributions. The policies that do the most to improve workers' wages by increasing the capital–labor ratio also drive down the return on capital. Policies that appear attractive because they enhance or at least preserve the rate of return on pension contributions appear in a less favorable light if we consider their effect on workers' total lifetime income.

APPENDIX 5A

Table 5A.1 Internal rates of return under social security reform plans, by earnings profile in a static simulation

Year of retirement	Earnings profile									
	Low wage			Average wage			High wage			Composite
	1	2	3	4	5	6	7	8	9	
Policy option 1: Pay-go tax increase										
2000	5.7	5.9	6.7	4.0	4.3	5.0	3.4	3.1	3.5	4.0
2025	4.6	4.9	5.5	3.2	3.4	4.2	2.3	2.2	2.7	3.2
2050	4.4	4.5	5.1	2.8	2.9	3.4	1.9	1.7	2.0	2.7
2075	3.9	4.0	4.6	2.3	2.5	3.0	1.5	1.3	1.7	2.3
Policy option 2: Pre-funding tax increase										
2000	5.7	5.9	6.7	4.0	4.3	5.0	3.4	3.1	3.5	4.0
2025	4.4	4.6	5.2	2.9	3.0	3.6	2.0	1.8	2.3	2.8
2050	4.2	4.3	4.9	2.6	2.8	3.3	1.7	1.6	2.0	2.6
2075	4.0	4.1	4.7	2.5	2.6	3.2	1.6	1.5	1.9	2.4
Policy option 3: Pay-go OASDI benefit cut + DC pension										
2000	5.6	5.9	6.7	4.0	4.2	5.0	3.4	3.1	3.5	4.0
2025	4.0	4.1	4.7	2.7	2.9	3.3	2.0	1.9	2.3	2.7
2050	4.0	4.1	4.5	3.0	3.0	3.3	2.5	2.3	2.5	2.9
2075	3.9	4.0	4.4	2.9	2.9	3.2	2.4	2.2	2.4	2.8

Social Security pension *only*										
2000	5.6	5.9	6.7	4.0	4.2	5.0	3.4	3.1	3.5	4.0
2025	3.9	4.1	4.7	2.4	2.6	3.2	1.5	1.4	1.8	2.4
2050	3.8	3.9	4.5	2.2	2.4	2.9	1.4	1.2	1.6	2.2
2075	3.6	3.8	4.4	2.1	2.2	2.8	1.2	1.1	1.5	2.0
Level annuity *only*										
2000	–	–	–	–	–	–	–	–	–	–
2025	4.3	4.1	3.8	4.2	4.1	3.9	4.2	4.1	4.0	4.1
2050	4.6	4.5	4.4	4.5	4.4	4.2	4.5	4.4	4.3	4.4
2075	4.6	4.5	4.4	4.5	4.4	4.2	4.4	4.4	4.3	4.4

Note: Earnings profiles 1, 4 and 7 show declining earnings over a career; profiles 2, 5 and 8 have level earnings; and profiles 3, 6 and 9 have rising earnings. The analysis is restricted to workers who actually reach age 62 and it uses a common projection of expected survival rates.

Source: Authors' tabulations.

Table 5A.2 *Change in lifetime earnings, pensions and net income*[a] *under alternative pension reforms, by lifetime earnings profile and retiree cohort, static simulation*

Year of retirement	Earnings profile									
	Low wage			Average wage			High wage			Composite
	1	2	3	4	5	6	7	8	9	
Policy option 1: Pay-go tax increase										
Lifetime earnings										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2050	98.8	98.2	97.8	98.4	97.9	97.6	98.2	98.0	97.7	98.0
2075	97.0	96.7	96.5	96.9	96.7	96.5	96.8	96.7	96.6	96.7
Lifetime pension										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2050	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2075	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Lifetime income[b]										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2050	99.2	98.7	98.5	98.7	98.3	98.1	98.4	98.3	98.1	98.4
2075	98.0	97.7	97.6	97.5	97.3	97.2	97.3	97.2	97.1	97.3

Policy option 2: Pre-funding tax increase										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	99.3	98.9	98.7	99.0	98.8	98.5	98.9	98.8	98.6	98.8
2050	98.1	97.9	97.8	98.0	97.9	97.8	98.0	97.9	97.8	97.9
2075	97.5	97.3	97.1	97.4	97.3	97.1	97.3	97.3	97.2	97.3
Lifetime pension										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2050	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2075	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
Lifetime income[b]										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	99.5	99.2	99.1	99.2	99.0	98.8	99.1	98.9	98.8	99.0
2050	98.7	98.6	98.5	98.4	98.3	98.2	98.3	98.2	98.2	98.3
2075	98.3	98.1	98.0	98.0	97.8	97.7	97.8	97.7	97.6	97.8
Option 3: Pay-go benefit cut + 2.6% DC pension										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	98.8	98.2	97.9	98.4	98.0	97.6	98.2	98.0	97.8	98.1
2050	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3
2075	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3	97.3
Lifetime social security pension										
2000	97.9	97.9	97.9	97.9	97.9	97.9	97.9	97.9	97.9	97.9
2025	76.6	76.6	76.6	76.6	76.6	76.6	76.6	76.6	76.6	76.6
2050	70.6	70.6	70.6	70.6	70.6	70.6	70.6	70.6	70.6	70.6
2075	67.5	67.5	67.5	67.5	67.5	67.5	67.5	67.5	67.5	67.5

Table 5A.2 (*continued*)

Year of retirement	Earnings profile									
	Low wage			Average wage			High wage			Composite
	1	2	3	4	5	6	7	8	9	
Lifetime individual account pension										
2000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2025	8.9	12.3	11.1	21.5	24.1	23.3	31.6	34.0	32.2	24.7
2050	32.0	28.6	22.5	51.5	43.3	32.8	64.3	62.2	49.9	47.3
2075	30.8	27.4	21.6	49.5	41.6	31.5	61.7	59.8	48.0	45.4
Lifetime income[b]										
2000	99.4	99.4	99.4	99.6	99.6	99.6	99.7	99.7	99.7	99.6
2025	94.6	95.5	94.9	98.3	98.5	98.1	99.7	99.7	99.4	98.6
2050	99.1	97.9	96.0	102.4	100.6	98.5	103.1	102.4	100.9	101.2
2075	97.6	96.6	94.7	101.4	99.7	97.7	102.4	101.8	100.2	100.3

Notes: Earnings profiles 1, 4 and 7 show declining earnings over a career; profiles 2, 5 and 8 have level earnings; and profiles 3, 6 and 9 have rising earnings. Additions to Trust Fund reserves do not add to national saving.

[a] Net earnings as a percentage of the baseline value; pensions as a percentage of OASDI pensions in baseline; and lifetime net income (net earnings plus pensions) as a percentage of value in baseline.

[b] Lifetime income is the sum of lifetime earnings and lifetime pensions from the reformed pension system. The change is reported for individuals retiring in the years shown in column 1.

Source: Authors' tabulations.

Table 5A.3 Change in lifetime earnings, pensioins and net income[a] under alternative pension reforms, by lifetime earnings profile and retiree cohort, dynamic simulation

Year of Retirement	Earnings profile									
	Low wage			Average wage			High wage			Composite
	1	2	3	4	5	6	7	8	9	
Policy option 1: Pay-go tax increase										
Lifetime earnings										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.5	101.0	101.3	100.7	101.1	101.4	100.9	101.1	101.3	101.0
2050	100.2	99.4	99.0	99.6	99.1	98.6	99.3	99.1	98.8	99.1
2075	98.2	98.0	97.8	98.1	98.0	97.9	98.1	98.0	97.9	98.0
Lifetime pension										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.7	100.7	100.7	100.7	100.7	100.7	100.7	100.7	100.7	100.7
2050	101.4	101.4	101.4	101.4	101.4	101.4	101.4	101.4	101.4	101.4
2075	101.7	101.7	101.7	101.7	101.7	101.7	101.7	101.7	101.7	101.7
Lifetime income[b]										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.5	100.9	101.1	100.7	101.0	101.3	100.9	101.0	101.2	101.0
2050	100.6	100.0	99.7	100.0	99.5	99.2	99.6	99.4	99.2	99.6
2075	99.4	99.2	99.1	98.9	98.8	98.7	98.7	98.6	98.5	98.7

Table 5A.3 (continued)

Year of	Earnings profile									
	Low wage			Average wage			High wage			Composite
Retirement	1	2	3	4	5	6	7	8	9	
Policy option 2: Pre-funding tax increase										
Lifetime earnings										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.9	102.9	104.9	101.7	103.3	104.9	102.4	103.3	104.2	103.1
2050	108.3	108.9	109.2	108.9	109.4	109.7	109.2	109.4	109.6	109.3
2075	110.5	110.5	110.3	110.6	110.5	110.5	110.6	110.5	110.5	110.5
Lifetime pension										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	111.6	111.6	111.6	111.6	111.6	111.6	111.6	111.6	111.6	111.6
2050	114.0	114.0	114.0	114.0	114.0	114.0	114.0	114.0	114.0	114.0
2075	114.1	114.1	114.1	114.1	114.1	114.1	114.1	114.1	114.1	114.1
Lifetime income[b]										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	104.2	105.4	106.8	103.6	104.9	106.2	103.7	104.5	105.3	104.6
2050	110.2	110.4	110.7	109.9	110.3	110.6	109.9	110.0	110.3	110.2
2075	111.8	111.6	111.5	111.4	111.3	111.2	111.2	111.1	111.1	111.2

Policy option 3: Pay-go OASDI benefit cut + DC pension										
2000	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2025	100.4	102.2	104.1	101.1	102.6	104.0	101.7	102.6	103.4	102.4
2050	107.6	108.7	109.5	108.4	109.3	110.0	108.8	109.3	109.7	109.2
2075	111.7	111.9	112.1	111.9	112.1	112.3	112.0	112.1	112.2	112.1
Lifetime social security pension										
2000	99.0	99.0	99.0	99.0	99.0	99.0	99.0	99.0	99.0	99.0
2025	91.4	91.4	91.4	91.4	91.4	91.4	91.4	91.4	91.4	91.4
2050	82.3	82.3	82.3	82.3	82.3	82.3	82.3	82.3	82.3	82.3
2075	78.1	78.1	78.1	78.1	78.1	78.1	78.1	78.1	78.1	78.1
Lifetime individual account pension										
2000	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
2025	7.2	10.4	9.8	17.6	20.4	20.2	26.2	28.7	27.6	20.8
2050	18.9	18.0	15.0	32.0	28.8	23.5	41.2	41.3	34.6	31.0
2075	16.2	15.7	13.3	27.9	25.5	21.2	36.2	36.6	31.1	27.4
Lifetime income[b]										
2000	99.7	99.7	99.7	99.8	99.8	99.8	99.9	99.9	99.9	99.8
2025	99.8	102.1	103.3	102.6	104.3	105.5	104.0	105.0	105.7	104.2
2050	105.5	106.2	105.8	109.6	109.6	109.2	111.1	111.4	110.9	109.9
2075	105.8	106.3	105.5	110.7	110.4	109.5	112.4	112.5	111.7	110.8

Notes: Earnings profiles 1, 4 and 7 show declining earnings over a career; profiles 2, 5 and 8 have level earnings; and profiles 3, 6 and 9 have rising earnings. Additions to Trust Fund reserves do not add to national saving and are invested in the United States.

[a] Net earnings as a percentage of the baseline value; pensions as a percentage of OASDI pensions in the baseline; and lifetime net income (net earnings plus pensions) as a percentage of value in the baseline.

[b] Lifetime income is the sum of lifetime earnings and lifetime pensions from the reformed pension system. The change is reported for individuals retiring in the years shown in column 1.

Source: Authors' tabulations.

NOTES

* Prepared for the International Forum of Collaboration Projects meetings in Osaka, Japan, 22–24 September 2000, sponsored by the Economic and Social Research Institute (ESRI). An earlier version of this chapter was prepared for the Joint Conference for the Retirement Research Consortium, 'The Outlook for Retirement Income', held on 17–18 May 2000, in Washington, DC. Partial support for the research described here was obtained under grants from ESRI to Brookings and from the US Social Security Administration (SSA) to the Center for Retirement Research at Boston College. The opinions and conclusions are solely those of the authors and do not represent those of the ESRI, the Social Security Administration, the Center for Retirement Research or the Brookings Institution.

1. Aaron et al. (1989); Bosworth and Burtless (1997).
2. Further details about the data set are provided in Toder et al. (1999, especially chs 2 and 8). See also Bosworth et al. (2000) and Bosworth and Burtless (2000).
3. Our classification scheme ignores a worker's earnings before age 32, because nearly all workers have low but sharply rising earnings early in their careers. Many workers have very low earnings while they are in their twenties because they are still studying. The cutoff values for classifying workers into one of the nine possible groups were originally derived from actual and projected earnings of individuals born between 1931 and 1960 (Toder et al. 1999, ch 8). While workers in the 1931–40 birth cohort are evenly distributed between the three groups for purposes of measuring their average earnings level, the use of a classification based on the larger population resulted in a somewhat disproportionate number of workers being allocated to the categories of rising and declining earnings. Note that our estimates of career-average earnings include years in which individual workers may have zero earnings. When the SSA calculates economy-wide average earnings in a given calendar year, workers with zero earnings are obviously excluded from the calculation.
4. We computed the trend in career earnings as $t=(C-A)/(C+A)$, where A is the worker's earnings between ages 32 and 41 and C is average earnings between ages 52 and 61. When $t<-1/9$ we classified the career earnings profile as 'declining;' when $t>1/9$ we classified the profile as 'rising.' Other career earnings paths were treated as 'level'. Note that a worker's earnings each year were calculated as the ratio of his or her actual earnings to the economy-wide average earnings in that year.
5. The low, average and high earnings patterns assumed by the US SSA are based on the assumption that the three classes of workers earn, respectively, 45 percent of the economy-wide average wage in every year, the average wage, and the maximum taxable wage.
6. We used the social security actuary's intermediate mortality projections to estimate future unisex life tables for all workers who survive to age 62. These life tables imply continuing improvements in age-specific survival rates over the next century.
7. The Trustees' most recent projections reflect some improvement in the social security outlook. Under the intermediate assumptions of the 2003 Trustees' Report, the OASDI Trust Fund will not be depleted until 2042 (Board of Trustees, OASDI 2003). However, the changes in the Trustees' assumptions since 1996 have little effect on our simulation results. They have absolutely no effect on the qualitative conclusions drawn in this chapter.
8. The OASDI 'cost rate' is the sum of annual benefit and program administration costs divided by the taxable payroll. The projected cost rate in 2000 is 11.8 percent. If we allowed the payroll tax rate to decline, under a pure pay-as-you-go rule the tax rate in 2000 would obviously fall below its current rate, 12.4 percent. Policymakers who are intent on preserving the current benefit structure are unlikely to permit payroll taxes to fall in the face of large anticipated increases in the OASDI cost rate. For that reason, we do not think it is interesting to consider the impact of a pure pay-go tax rate.
9. This option is the same as that proposed in Aaron et al. (1989). It avoids many of the

transitional problems of a shift away from a pay-go system by seeking only to pre-fund additions to the current cost burden. The 75-year planning horizon is viewed as representative of the remaining life expectancy of young contributors. If we had implemented the solvency test using a planning horizon of 25 years or less, the initial tax increase in 2000 and the Trust Fund buildup would be much smaller.

10. Recall that our simulation is calibrated to the assumptions in the 1996 Trustees' Report. That report showed an actuarial imbalance in 1996 equal to 2.2 percent of taxable payroll. For the 75-year planning period beginning in 2000, the imbalance had increased to 2.6 percent under the 1996 assumptions because of the addition of four deficit years to the planning period. The 2.6 percent imbalance is larger than the one shown in the Trustees' Report for 2003, which shows a 75-year imbalance of 1.9 percent. The difference is mainly explained by changes in economic assumptions between 1996 and 2003. Much of the improvement in the economic outlook can be traced to changes in the way price change is measured by the consumer price index (CPI).
11. For purposes of this simulation, we assume that the benefit cuts are imposed equally on all recipients collecting social security pensions, including pensioners already on the rolls. Our model also permits us to calculate the reduction in benefits that would be needed if only *new* pensioners were affected by a benefit cut. Under a myopic pay-go rule the benefit cuts would begin around the same time and proceed more abruptly than under the assumption that all current beneficiaries shared the benefit cut. If instead the benefit cuts were phased in to preserve pay-go solvency and benefit cuts only imposed on new beneficiaries, the benefit cuts would begin earlier and proceed more gradually, but the ultimate reduction in benefits would not differ much from that shown in the lower panel of Figure 5.2.
12. The replacement rate nonetheless offers a very direct measure of pension adequacy. Comparisons across workers can highlight the redistributional impact of the current system. In our baseline case, which assumes the maintenance of the current benefit formula, the replacement rate ranges from a high of about 80 percent for low-wage workers to 30 percent for high-wage workers.
13. Our annuity calculations assume that insurance companies selling annuities expect to earn the riskless (Treasury bond) rate of return on their investments.
14. This difference between women and men was more pronounced among workers born between 1931 and 1940 than it will be in the future. Women born since 1950 have been less likely than women in earlier generations to leave the labor force for long periods after the birth of their first child. Even today, however, many women have depressed earnings in years when they are rearing young children, and this tends to produce a career earnings profile in which wages are relatively high late in the worker's career.
15. We observe that large numbers of workers accumulate few or no financial assets by the time they retire. Economists do not know, however, whether the lack of savings implies that workers have substituted the promise of social security benefits for private wealth accumulation or if workers are simply myopic.
16. If R_t is the level of nominal reserves at the end of year t and Δp is the change in the price level between $t-1$ and t, then the change in real reserves is $R_t - R_{t-1}(1+\Delta p)$.
17. The structure of the macroeconomic model is discussed more fully in Bosworth and Burtless (1997, 2000).
18. We performed additional simulations to test the sensitivity of our results to this assumption. We allowed most of the extra saving to be invested outside the United States. The major difference is that there is no change in the domestic capital–output ratio and hence no decline in the rate of return. On the other hand, domestic wages are also unchanged. With respect to the retirement accounts, the assumption of investing abroad produces results very similar to our static case.
19. The reduction of the bond rate does have a feedback effect on the calculation of social security actuarial balance for option 2. The loss of interest earnings offsets part of the improvement in OASDI finances that would otherwise result from the gain in tax revenues.

20. By reducing the relative importance of the DC pension, the dynamic simulation also eliminates much of the distinction between workers with rising and declining wage profiles.

REFERENCES

Aaron, Henry J., Barry P. Bosworth and Gary Burtless (1989), *Can America Afford to Grow Old? Paying for Social Security*, Washington, DC: Brookings Institution.

Board of Trustees, Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds (OASDI) (1996), *1996 Annual Report of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds,* Washington, DC: Social Security Administration.

Board of Trustees, Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds (OASDI) (2003), *2003 Annual Report of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds.* Washington, DC: Social Security Administration.

Bosworth, Barry P., and Gary Burtless (1997), 'Social security reform in a global context', in Steven A. Sass and Robert K. Triest (eds), *Social Security Reform: Links to Saving, Investment, and Growth*, Boston, MA: Federal Reserve Bank of Boston, pp. 243–74.

Bosworth, Barry P., and Gary Burtless (2000), 'The effects of social security reform on saving, investment, and the level and distribution of worker well-being', Working Paper 2000–02, Chestnut Hill, MA: Boston College Center for Retirement Research.

Bosworth, Barry P., Gary Burtless and Eugene Steuerle (2000), 'Lifetime earnings patterns, the distribution of future social security benefits, and the impact of pension reform', *Social Security Bulletin*, **63** (4) 74–98.

Toder, Eric, Cori Uccello, John O'Hara, Melissa Favreault, Caroline Ratcliffe, Karen Smith, Barry Bosworth and Gary Burtless (1999), *Modeling Income in the Near Term – Projections of Retirement Income Through 2020 for the 1931–60 Birth Cohorts*, Washington, DC: Urban Institute.

6. Asset accumulation and retirement income under individual retirement accounts: evidence from five countries

Gary Burtless*

1. INTRODUCTION

As populations in rich countries grow older, the cost of paying for public pensions has risen, boosting tax burdens and placing increased pressure on government budgets. Only one of the seven largest industrial countries, the United Kingdom, has overhauled its public pensions in a way that is likely to hold down future pension spending so that it does not increase sharply relative to national income. The favorable outlook for public spending on British pensions is the result of policies that tightly restrain the growth of basic government pensions and encourage active workers to abandon the second-tier, earnings-related public program in favor of private pensions. Future retirees are expected to derive much more of their retirement income from privately managed and invested pension accounts than from publicly financed, pay-as-you-go pensions. Other leading industrial countries still face major challenges in paying for or fundamentally reforming their main public pension programs (Bosworth and Burtless 1998).

Policymakers in several rich countries show interest in following the British example and replacing part of their public systems with private pensions organized around individual retirement accounts. In May 2001 the German government revised Germany's national pension system to curtail the future growth of publicly provided pensions and to subsidize the creation of new defined-contribution pensions based upon individual accounts. In June 2001 the upper house of the Japanese legislature gave final approval to the government's plan to offer workers tax-favored retirement saving plans modeled closely on 401(k) retirement accounts now available in the United States. The new retirement plan, like the one in Germany, is intended to supplement pensions provided by the main public system. In both Germany and Japan, benefits under the main public system will be scaled back for workers retiring over the next several decades. The United States has long used tax incentives to promote private retirement systems, which now cover about half of the workforce. Many critics of traditional public pensions would like to go much further. A presidential commission recently outlined three reform plans to reduce benefits under the existing US social security system and replace them with annuities financed out of voluntary retirement savings accounts (President's Commission to Strengthen Social Security 2002).

This chapter examines evidence on the likely success of individual retirement accounts in providing retirement incomes to typical workers. Historical and simulated data on financial market performance are used to evaluate the market risks facing contributors to a private system based on individual retirement accounts. The chapter provides evidence on these risks by considering the hypothetical pensions that workers in five industrialized countries would have received based on financial market performance between 1927 and 2002 if they had accumulated retirement savings in individual accounts. The contributors to individual retirement accounts are assumed to have identical careers and to contribute a fixed percentage of their wages to private investment funds. When contributors reach retirement age, they convert their retirement savings into a level annuity. To make the calculations comparable across countries and across time, all contributors are assumed to have an identical career path of earnings and to face the same mortality risks after reaching retirement. Contributors differ only with respect to the level and timing of stock and bond returns, bond yields when they reach retirement, and price inflation. These differences occur because of the differing start and end dates of the workers' careers and because workers reside in different countries and are assumed to restrict their investments to the stocks and bonds of their own country.

The analysis demonstrates that the financial market risks of a funded private retirement system are empirically large in all of the industrialized countries. Although some of these risks are also present in a public retirement system, a public system, backed by the taxing and borrowing authority of the state, can spread risks over a much larger population of potential contributors and beneficiaries. This makes the risks more manageable for individual workers, many of whom have little ability to insure themselves privately against financial market risk.

2. RISK AND RETURN IN AN INDIVIDUAL ACCOUNT SYSTEM

The main goal of a pension system is to replace labor earnings that are lost as a result of old age, premature death and invalidity. The usual way rich countries achieve this goal is through mandatory, publicly financed pensions. The typical public program is a defined-benefit program in which the pension is calculated on the basis of the worker's years of coverage under the system and average covered wages while the worker is contributing to the system (World Bank 1994, esp. pp. 102–9). Benefits are largely financed out of current contributions of active workers and their employers. Only a few public systems have built up enough reserves to pay for a large percentage of future pension obligations. Excessive pension commitments and a growing ratio of retired to active workers have pushed many pay-as-you-go pension programs towards insolvency. Governments can restore solvency through higher taxes, reduced monthly pensions, or a delay in the age at which workers can claim benefits. Many industrialized countries, including the United States, have taken one or more of these steps, but their pension systems still face major funding shortfalls.

Private retirement accounts, in contrast, are usually operated as funded programs. Moreover, many advocates of reform believe that a new retirement system should be built around defined-contribution rather than defined-benefit pensions. The US employer-sponsored pension system has already seen a major shift toward defined-contribution plans, which now cover two-thirds of active participants and own more than 50 percent of the assets held by the private pension system (EBRI 2002). Instead of contributing to a collective, pay-as-you-go retirement program, workers in defined-contribution plans build up retirement savings in individually owned and directed private accounts. Workers can withdraw their funds from the accounts when they become disabled or reach the retirement age, and their heirs can inherit any funds accumulated in the account if the worker dies before becoming disabled or reaching the retirement age. At the time a worker chooses to start receiving a pension, some or all of the funds in the worker's account are converted into an annuity that lasts until the worker dies. In most plans, workers are free to decide how their contributions are to be invested, at least within broad limits.

Private defined-contribution pension plans differ from public systems in two important ways. First, there is little or no redistribution of benefits from high-wage to low-wage workers. The ultimate retirement benefit depends on each worker's own contributions and the success of the worker's investment plan. Workers who make larger contributions receive bigger pensions, other things equal; workers whose investments earn high returns enjoy higher retirement incomes than workers whose savings earn poor returns.

Second, in a private system, workers' pensions are paid out of accumulations of their own previous savings. In contrast, public pensions are financed mainly by the payroll taxes of active workers and their employers. This difference between the two kinds of system implies that the savings accumulation in a private plan would be many times larger than the reserves needed in a pay-as-you-go public system. For this reason, existing pay-as-you-go systems cannot simply be scrapped in favor of an individual account system. Workers who are currently retired or near retirement have already made contributions that entitle them to future pensions. The pension liabilities of the old system must be financed out of future contributions, new taxes or public borrowing, rather than out of funds that have already been accumulated. This represents a major hurdle to the introduction of a new funded retirement system. One or more generations of active workers will be required to pay for the liabilities of the existing system regardless of whether most of their own pensions will be financed out of that system.

This chapter focuses on workers' returns on their contributions to individual retirement accounts rather than on their returns from the overall pension system. Readers should bear in mind, however, that returns obtained under the existing public system and under an individual retirement account are not strictly comparable. Contributions to existing public programs include a large implicit tax to pay for the unfunded liabilities that were accumulated in paying the benefits of deceased and already-retired workers who did not fully contribute to the system. Virtually all of this tax will have to be paid irrespective of whether the present public system is maintained or is replaced with a new system of private accounts. To make a meaningful comparison between the contribution rates to public and to individual-account systems, it is necessary either to subtract this implicit tax from the public pension contribution rate or to add it to the rate needed to fund the new private accounts.[1]

3. FINANCIAL RISK

A defined-contribution system allocates risks in a very different way from a collective, defined-benefit system. Under most public pension systems, workers born in the same year who have a similar earnings record and have the same number of dependents receive similar retirement benefits. Because of political constraints on legislators, the public pension formula changes very slowly and only after protracted political debate. Since this debate involves both contributors and beneficiaries, changes in contribution and benefit formulas tend to reflect a compromise between the interests of the two groups. The effects of unanticipated demographic, labor market and financial market developments are rarely if ever borne by a single cohort. They are spread across a number of cohorts through gradual changes in contribution rates and benefit levels. In contrast, workers participating in a defined-contribution system bear many more of the risks associated with financial market fluctuations.

Workers enrolled in a defined-contribution pension plan face three kinds of financial market risk. They are exposed to the risk that the real return on their contributions will fall short of the historical average over the course of their working career. If workers obtain unexpectedly low returns on their retirement savings, they will enter old age with insufficient savings to finance a comfortable retirement.

Second, at the point they retire workers may find it expensive to purchase annuities. Workers who want to ensure that they will not outlive their assets will convert some or all of their retirement savings into an annuity around the time they retire. The market price they pay for annuities depends on four factors: their expected life span when they purchase annuities, the amount of adverse selection among the population buying annuities, the profit requirements needed to induce insurance companies to offer annuities, and the expected return on assets in which insurance companies will invest their reserves. Even assuming that mortality risk among workers of the same age is identical, that adverse selection among potential annuity buyers is negligible, and that insurance companies will sell annuities at zero profit, workers will still pay widely varying prices for annuities over time because of fluctuations in expected returns on insurance company reserves.

Finally, workers who buy nominal annuities are subject to inflation risk. Inflation that occurs after a worker retires can have a dramatic impact on the purchasing power of a pension. If inflation turns out to be unexpectedly high, retired workers may reach advanced old age with a pension that has little purchasing power. This risk could be avoided if workers purchased price-indexed annuities rather than level nominal annuities. Indexed annuities are now available in the United Kingdom, though not in most other industrial countries (Brown et al. 2001). In countries where indexed annuities are not available, retired workers have no obvious way to ensure stable consumption after retirement.

4. HISTORICAL RETURNS IN THE UNITED STATES

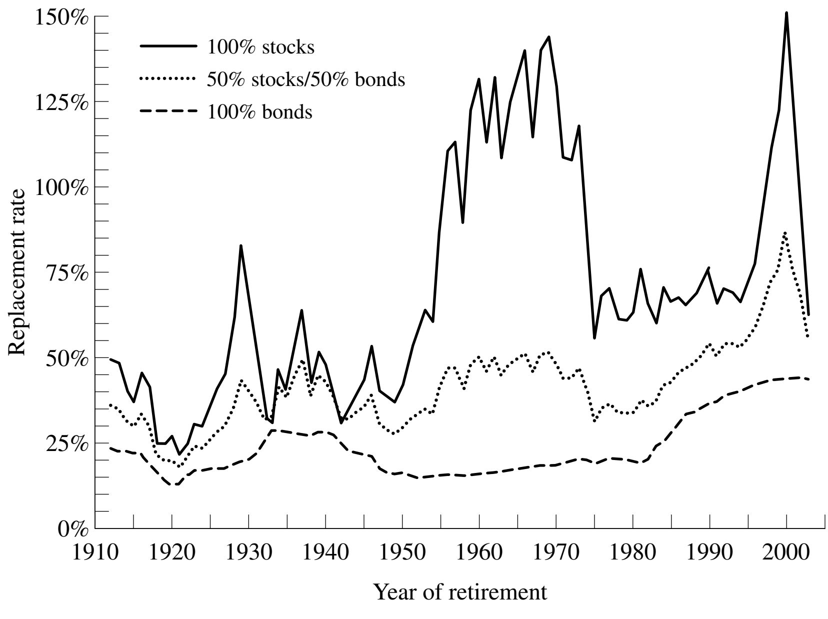
The risks just mentioned are relevant to considering whether an individual account pension system can deliver dependable income replacement in old age. One way to evaluate these risks is to calculate the real pensions workers would have obtained if they had contributed to a defined-contribution plan in the past. To calculate such pensions it is necessary to define a standard career path of earnings and pension contributions, calculate the assets that would be accumulated under a chosen investment strategy, and estimate the real pension the worker could purchase with the assets accumulated at retirement. All the calculations that follow are based on the earnings profiles of workers who have a full, 40-year career that begins at age 22 and ends at 62. In the absence of economy-wide wage growth, workers are assumed to have a lifetime path of real earnings that matches the age–earnings profile of employed US men in 1995 (US Census Bureau 1996, p. 34). In 1995 the earnings of 22-year-old American males were roughly one-quarter those of 45-year-olds, while earnings of 60-year-olds were 17 percent less than those of 45-year-olds. The career path of earnings is also affected by the growth of real wages in the wider economy, which for purposes of this exercise is assumed to be 1.5 percent a year, approximately the growth rate of US real wages since the Second World War.

I calculate the value of savings at retirement using two main assumptions. Workers contribute to their pension plans on the first day of every year, and they follow a consistent investment strategy over their career. In particular, workers are assumed to invest their retirement savings in some desired combination of bonds and common stocks. All stock dividends during the year are reinvested in new stock purchases, and all bond interest payments are reinvested in a standard portfolio of long-term government bonds. If workers invest in a mixture of both stocks and bonds, they re-balance their portfolio at the end of each year to maintain the preferred allocation of stocks and bonds. The income flows from assets in the retirement savings account are assumed to be free of individual income taxes at the time they are reinvested. Unlike ordinary investors, who must pay trading fees and commissions when buying and selling most financial assets, workers are assumed to face no transaction costs in making these investments.

When workers attain age 62, they convert their accumulations into a single-life annuity that is fixed in nominal terms. The annuity seller bases its price on the expected mortality experience of American males who reached age 65 in 1995, using mortality projections of the Social Security Actuary (Board of Trustees, OASDI 2001). The Actuary's projections take account of gradual improvements in mortality experience that Americans are expected to enjoy over the next several decades. I assume the insurance company does not charge a load factor to cover its profit requirements or possible adverse selection among people seeking to buy annuities. Thus, retiring workers are assumed to purchase fair annuities.

In determining the sales price of an annuity, the insurance company assumes it will be able to invest the worker's funds at the long-term government bond yield prevailing when the annuity is purchased. Since the annual annuity payment is fixed in nominal terms, the insurance company uses the nominal bond yield in this calculation. Unlike most public pension programs, the insurance company does not adjust the nominal value of the annuity from year to year to reflect changes in the price level. Insurance companies in a few industrial countries can now offer indexed annuities, because they can purchase indexed government bonds. However, the historical experience with this kind of asset is too recent for us to calculate the price that would have been charged for real annuities in the past.

A common measure of the value of a pension is the replacement rate. Figure 6.1 shows replacement rates of hypothetical US workers who retired after 40-year careers that ended on 1 January of the years indicated along the horizontal axis. The workers are assumed to contribute 7 percent of their wages to the retirement savings account. The upper line shows



Note: Assumed contribution rate is 7% of wages.

Source: Author's tabulations of US equity and bond return data supplied by Global Financial Data (March 2003).

Figure 6.1 Individual retirement account replacement rates for US workers under alternative investment portfolios, 1912–2003 replacement rates in successive years for workers who invest in a portfolio of common stocks that earns the same total return as the Standard and Poor's composite stock index. The middle line shows replacement rates for workers who invest in a portfolio consisting one-half of US equities and one-half of long-maturity US government bonds. The lower line displays replacement rates obtained by workers who invest all of their retirement savings in long-maturity government bonds. The replacement rate is the worker's real annuity divided by his/her average real earnings between ages 54 and 58, when his/her lifetime earnings are at their peak. Replacement rates are measured at age 62, when workers first retire. For example, the first point along the top line shows the replacement rate of a worker who entered employment in 1872, contributed 7 percent of his/her earnings to a retirement account invested in US common stocks, and converted his/her retirement savings into a level annuity at the start of 1912.

For the 92 overlapping 40-year careers ending in 1912–2003, the average replacement rate based on a stock portfolio is slightly more than 70 percent, near the middle of the range recommended by financial planners as an income goal in retirement. However, the standard deviation of replacement rates is almost 34 percent, implying that the range of income replacement provided by stock-invested retirement savings plans is quite large. A worker who receives the ninth decile replacement rate would receive a pension that initially replaces 122 percent of his/her peak earnings, whereas a worker who receives the first decile replacement rate would collect a pension that replaces a little more than a quarter of this amount, just 33 percent of peak earnings. The range of replacement rates is reduced if workers steadily invest a higher percentage of their retirement savings in bonds or bills. This strategy substantially reduces the expected pension, but it may have only a slight positive impact on the first decile pension. For example, if 10 percent of retirement savings are invested in US government bonds, the expected replacement rate falls 7 percentage points, from 71 to 64 percent, but the first decile replacement rate increases only 0.3 percentage points, from 33.4 to 33.7 percent. A larger allocation to bonds actually reduces the first decile replacement rate. For example, with a 50–50 allocation to stocks and bonds the first decile replacement rate falls to just 28.6 percent. If the goal of a conservative investment strategy is to protect workers' pensions in very poor financial markets, the strategy of investing steadily in either long-maturity government bonds or bills offers poor protection against the risk of obtaining a very small pension.

Another way to evaluate the range of investment outcomes shown in Figure 6.1 is to ask how high a worker's contribution rate must be to provide a 70 percent replacement rate at age 62. The median contribution rate needed to achieve this goal if all savings are invested in the US stock market is 7.5 percent. The first and ninth decile contribution rates under this investment strategy are 4.0 and 14.7 percent, respectively. If workers steadily invest a fraction of retirement savings in bills or bonds, the wide range in required contribution rates can be cut, though of course the average required contribution would have to rise to offset the reduction in expected returns. For example, if workers maintained an investment portfolio in which half their savings are placed in US stocks and half in long-maturity government bonds, the median required contribution rate would rise to 12.1 percent, and the first and ninth decile contribution rates would increase to 9.0 and 17.1 percent, respectively. If all retirement savings were invested in long-maturity government bonds, the required contribution rates would increase still further to 12.0 and 30.7 percent, respectively. The wide range of required contribution rates, even under very conservative investment strategies, implies that workers face great uncertainty in deciding how much of their annual earnings to set aside in investment accounts.

Workers who purchase nominal annuities also face uncertainty about the value of their pensions in the years after they retire. The Social Security Administration mortality projections imply that three-quarters of 62-year-old men will survive to their 72nd birthday. The twentieth-century record of US inflation suggests that retirees should expect to receive real annuities at age 72 that are much smaller than the ones they received at age 62. We can use US inflation experience between 1912 and 2002 to calculate the decline in real pensions workers would have experienced if they retired between 1912 and 1993. For workers investing 100 percent of retirement savings in stocks, the average real replacement rate measured at age 72 is 19 percentage points (or one-quarter) lower than it is at age 62. In many cases the loss of purchasing power would have been much worse. Inflation in the 1970s and early 1980s was so high that the age-72 replacement rates of US workers retiring between 1975 and 1983 were approximately the same as those of workers who retired in the Great Depression.

5. CROSS-NATIONAL RESULTS

Data on inflation and financial returns can be used to predict pensions for countries in addition to the United States. Consistent data on the total returns of stocks and long-maturity government bonds are available for France, Germany, Japan, the United Kingdom and the United States for calendar years 1927–2002.[2] Table 6.1 shows real stock and bond returns in the five countries. Stock returns have substantially exceeded bond returns in all five countries. On average, the geometric mean equity return is 4.4 percentage points higher than the mean bond return, with the equity premium ranging from a low of 3.1 percent in the United Kingdom up to 5.2 percent in Japan. The standard deviation of real equity returns is approximately twice that of bond returns, implying that investors accept substantially more year-to-year risk when holding stocks rather than bonds. The risk of holding government bonds is far from trivial, however. For example, the standard deviation of annual returns on long-maturity Japanese bonds is approximately the same as the standard deviation of annual returns on US equities. Japanese bonds are less risky than Japanese equities, but they are not notably less risky than US equities.

Table 6.1 shows sizeable differences in real returns across countries. US investors have enjoyed the highest long-term returns on stocks, while Japanese investors have obtained the lowest returns (see also Dimson et al, 2002). One dollar invested in the US stock market at the end of 1926 would have returned \$151 by the end of 2002. One yen invested in the Japanese stock market at the end of 1926 would have returned just 8.9 yen at the end of 2002. Stock market performance in the three European countries, especially in the United Kingdom, has been closer to US than to Japanese experience. Investors in all countries have experienced periods in which equity returns were persistently above or below average. The persistence of equity returns is especially notable in Japan. Japanese investors enjoyed an extraordinary 15 percent annualized rate of return on equities between 1948 and 1989, but this outstanding performance was counterbalanced by –12 percent annual returns between 1927 and 1947 and –9 percent returns after 1989. Japanese investors have also obtained relatively poor long-term returns on their bond investments, though these returns improved substan-

Table 6.1 Annual real investment returns in five industrial countries, 1927–2002 (%)

Asset class	France	Germany	Japan	UK	USA
Stocks					
Geometric mean	4.2	5.0	2.9	5.9	6.8
Arithmetic mean	8.1	9.2	8.3	8.1	8.9
Standard deviation	31.0	29.4	31.9	21.7	20.8
Long government bonds					
Geometric mean	−0.9	1.1	−2.3	2.8	2.2
Arithmetic mean	0.5	3.7	0.6	3.1	2.6
Standard deviation	15.9	15.1	20.1	8.9	9.6

Source: Authors' tabulations of total return and inflation statistics provided by Global Financial Data (updated March 2003).

tially after the mid-1970s. UK and US investors have earned the highest returns on bond investments, primarily because their governments have not defaulted on the public debt through high inflation or major currency reform. Much of the variability in French, German and Japanese bond returns can be traced to high inflation and currency reform in the immediate post-war period, when outstanding government bonds lost most of their value.

The large difference in average stock and bond returns means that workers who invest in stocks ordinarily accumulate more pension assets than workers who invest in bonds or in a combination of stocks and bonds. US workers retiring after a 40-year career ending in 1967 through 2003, for example, would on average have accumulated almost three times more assets if they had invested exclusively in stocks compared with their accumulation if they had invested exclusively in bonds. In the same period Japanese workers would have accumulated four times more assets with stock investment than bond investment.

Stock market investments delivered widely varying accumulations over time, however. A Japanese worker who invested solely in stocks and retired at the beginning of 1973 would have accumulated almost 11 times as much retirement savings as a worker who followed the same investment strategy and retired at the beginning of 2003. The gap between the accumulation of the luckiest and the unluckiest saver is not as large in Europe or the United States, but in all five countries the gap is big enough so that it would produce dramatic differences in workers' initial replacement rates.

Interestingly, both stock and bond returns have tended to converge in the five countries, probably as a result of the closer integration of world capital markets. The gap between the highest and lowest government bond return was just 1.5 percentage points in 1974–2002, whereas it was more than 6 points in the early post-war period and 13 points in the period from 1927 to 1946. Similarly, the difference between the best and worst stock market performance narrowed dramatically. The cross-country gap between the best and worst annualized equity return was more than 15 percentage points in 1927–46 but less than 5 percentage points in 1974–2002.

The observation period contains a total of 76 years, so it is straightforward to predict the pensions of 37 workers, namely, those who start their retirement at the beginning of successive years from 1967 to 2003. The pensions of these workers do not accurately reflect all the evidence on returns between 1927 and 2002, however, because annual returns at the beginning and the end of the 76-year period are reflected in the pensions of only one or two workers whereas returns in the middle of the 76 years are reflected in the pensions of essentially all of the simulated workers. To give equal weight to each annual observation, I created a sequence of annual returns so that each of the 76 annual observations is used an equal number of times.[3]

Table 6.2 shows selected information about simulated pensions based on cross-national data on stock market returns. The table contains information about pensions derived from two different investment strategies. In the top panel, workers are assumed to invest solely in equities throughout their careers. The bottom panel shows results for an investment strategy in which 60 percent of savings is allocated to equities and 40 percent is allocated to long-maturity government bonds, with annual re-balancing to maintain the

Table 6.2 Real returns and replacement rates under individual retirement accounts with alternative allocations to stocks and bonds (%)

	France	Germany	Japan	UK	USA
100% equities/0% bonds					
Required contribution rate	10.4	8.2	9.4	7.4	6.9
Internal rate of return					
Average	4.6	5.4	4.2	6.1	7.1
Standard deviation	3.0	2.8	8.2	1.7	1.7
Replacement rate at age 62					
Average	83	77	115	75	77
Ninth decile	137	134	247	118	118
Median	70	70	70	70	70
First decile	35	33	24	40	42
60% equities/40% bonds					
Required contribution rate	13.5	9.4	9.5	9.2	9.0
Internal rate of return					
Average	3.4	4.5	2.8	5.2	5.6
Standard deviation	3.9	2.5	7.7	1.6	1.6
Replacement rate at age 62					
Average	95	67	69	74	71
Ninth decile	176	90	123	104	96
Median	70	70	70	70	70
First decile	36	38	16	47	49

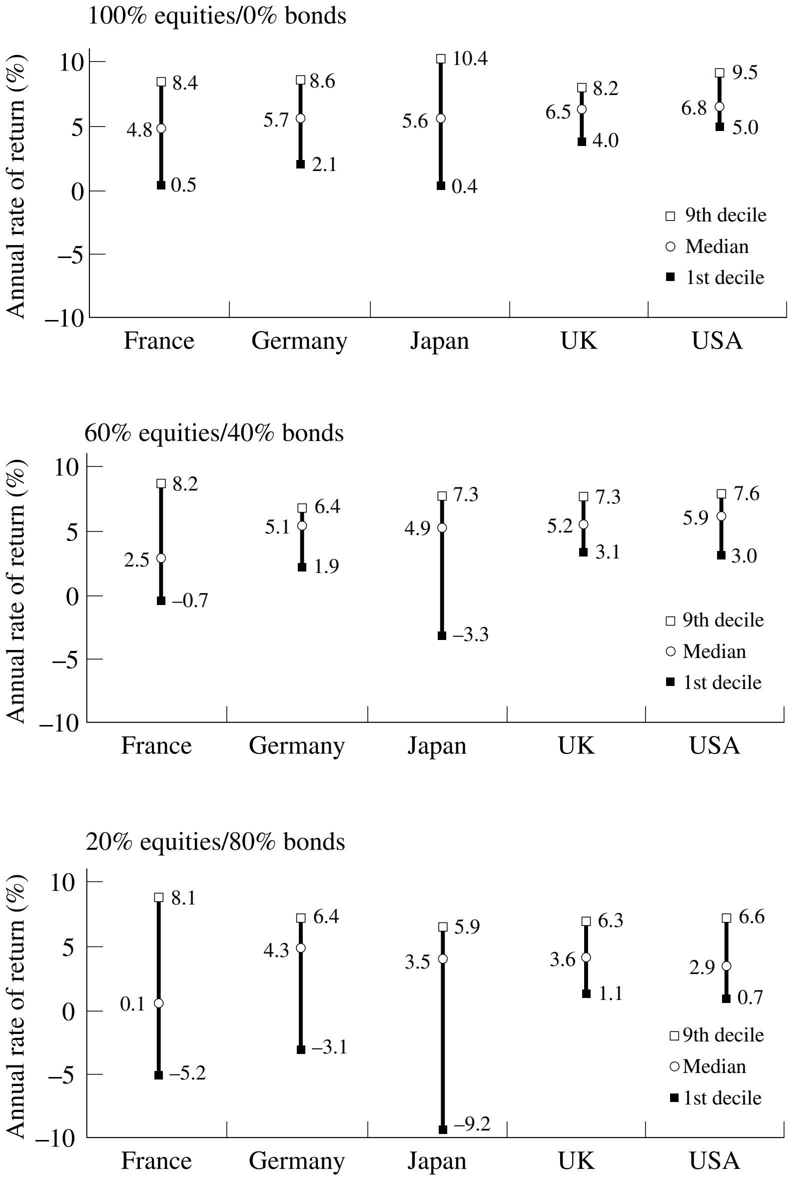
Source: Authors' tabulations as explained in text.

60/40 percent portfolio allocation. For each country and each investment strategy I have calculated pensions under the assumption that the worker's contribution rate is just high enough so that the *median* worker will have an initial pension that replaces 70 percent of his/her peak career earnings. The contribution rates needed to achieve this goal are displayed in the top row of each panel in the table.[4] For example, French workers who invest all their retirement savings in equities must contribute 10.4 percent of their wages to retirement saving if they want to have a 50 percent chance of obtaining a 70 percent (gross) replacement rate. German workers who invest exclusively in German stocks must contribute 8.2 percent of their wages to reach the same target.

When all retirement savings is invested in equities, US and UK workers face the lowest contribution rates while French and Japanese workers must make the largest contributions. The explanation for this difference is the lower return that French and Japanese workers can expect to earn on their stock investments. The second row in each panel of Table 6.2 shows the average real internal rate of return that workers obtain on contributions to their retirement saving accounts.[5] When retirement savings is invested exclusively in stocks, the cross-national pattern of returns is similar to the cross-national distribution of equity returns shown in Table 6.1. Not surprisingly, expected returns are uniformly lower for workers who invest part of their retirement savings in government bonds. For example, the mean return on an American worker's contributions falls from 7.1 percent to 5.6 percent as the percentage of savings allocated to bonds rises from 0 to 40. The decline in returns is somewhat smaller in the other countries. In every country except France the variability of returns shrinks when government bonds are included in the investment portfolio.

The effect of saving allocation on the distribution of investment returns is highlighted in Figure 6.2. The top panel in the graph shows the range of investment returns when retirement savings is invested exclusively in equities. In the lower two panels the allocation to bonds is progressively increased, first to 40 percent and then to 80 percent of retirement savings. For each country and investment allocation I have calculated the median return on retirement savings as well as the first and ninth decile returns. Based on investment experience between 1927 and 2002, a Japanese worker who invests solely in Japanese equities can expect a median real return on his/her retirement savings of 5.6 percent, roughly midway between the first decile return (0.4 percent) and the ninth decile return (10.4 percent). The gap between the first and ninth decile returns is smaller in France and Germany and dramatically narrower in the United States and the United Kingdom. Moreover, the median return on stock-invested savings is also higher in the United States and the United Kingdom than it is in Japan.

Perhaps surprisingly, the gap between the first and ninth decile returns tends to grow as a worker's allocation to bond investment rises, and this pattern is repeated in all five countries. In Germany, for example, the gap between the first and ninth decile returns is 6.5 percentage points for workers who invest all retirement savings in equities, whereas it is 9.5 percentage



Source: Authors' tabulations of inflation and equity and bond return data supplied by Global Financial Data (March 2003).

Figure 6.2 Distribution of real returns on pension contributions under alternative investment strategies points for workers who allocate 80 percent of their savings to government bonds. The gap increases because the first decile return falls steeply as more bonds are added to the investment portfolio. A larger allocation to bonds exposes workers to the risk that their retirement savings will be rapidly and permanently eroded as a result of unanticipated inflation or currency reform. Although equity investors are exposed to a similar risk, stock market shares represent an ownership claim on real assets. Even if the purchasing-power value of such claims falls in financial market crises, there is some prospect that the claims will regain some or all of their value when a crisis is over. The evidence of the past 76 years suggests that, compared with retirement savers who invest steadily in stocks, savers who invest in bonds are exposed to much greater risk of a very poor return. Furthermore, this risk is greater in France, Germany and Japan than it is in the United Kingdom or the United States. If the post-Second World War episodes of high inflation and currency reform are unlikely to be repeated, Figure 6.2 offers a misleading picture of the risks associated with bond investment.

Even when workers contribute the same percentage of their wages to retirement savings accounts and invest in the same portfolio of stocks and bonds, their pensions differ widely depending on the exact sequence of annual investment returns over their career. This range is displayed at the bottom of each panel in Table 6.2, which shows selected replacement rates for the five countries. The median replacement rate is identical across countries, because the contribution rate of each country was calibrated to ensure this result. The gap between the first and the ninth decile replacement rates differs widely across the five countries, however. When all retirement savings are invested in equities in the United States and United Kingdom, for example, the ratio of the ninth decile to the first decile pension ranges between 2.8 and 3.0. In France the ninth decile pension is 3.8 times the first decile pension, and in Japan the ratio is more than 10:1. Most workers would surely regard this range of uncertainty as excessive.

6. STRATEGIES TO DEAL WITH UNCERTAIN RETURNS

As demonstrated in Figure 6.1, replacement rates can vary enormously over short periods of time, even when workers invest their pension savings in a fairly conservative portfolio of equities and bonds. For example, a US worker who saved 7 percent of his/her salary and invested in a portfolio consisting of 50 percent stocks and 50 percent government bonds would have received a replacement rate of 87 percent if he/she retired in January 2000 but only 56 percent if he/she retired three years later in January 2003.

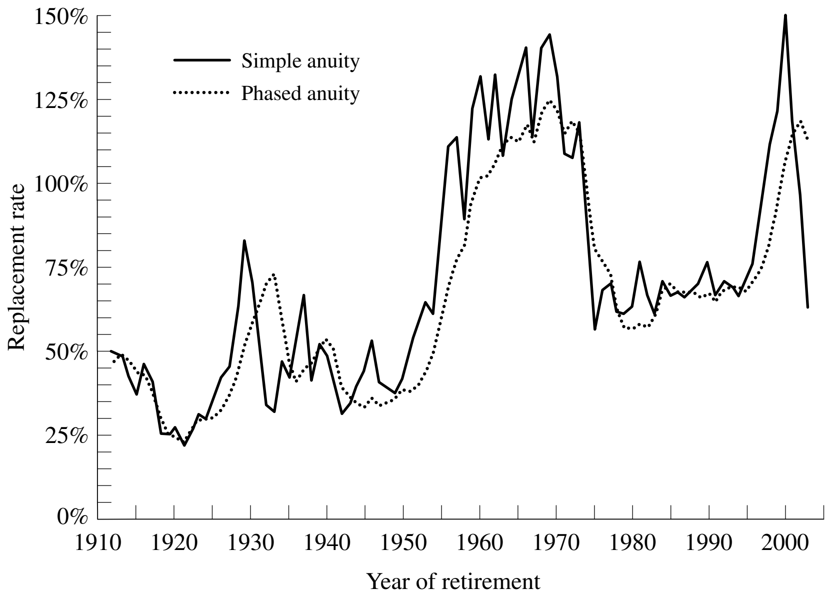
The value of pensions funded out of defined-contribution accounts depends critically on the interval in which workers accumulate stocks and bonds and the exact year when they convert their investment portfolio into annuities.

Workers can follow a couple of strategies to reduce the uncertainty of individual-account pensions. They can invest a larger portion of their retirement savings in safer assets, such as bonds or bills, rather than stocks. Replacement rates in the top and bottom panels of Table 6.2 can be compared to see how much the riskiness of pensions is reduced when workers reduce their equity allocation from 100 to 60 percent of retirement savings. One measure of the variability of pensions is the ratio of the ninth decile replacement rate to the first decile replacement rate. For German workers this ratio is 4:1 if all savings are invested in equities, but falls to 2.4:1 if 40 percent of retirement savings are invested in long-maturity government bonds. The ninth/first decile replacement rate ratio falls from 10.3 to 7.6 in Japan, from 3.0 to 2.2 in the United Kingdom, and from 2.8 to 2.0 in the United States as the percentage of savings allocated to bonds rises to 40 percent of the worker's portfolio. In France, a larger allocation to bonds actually increases the variability of replacement rates. On the whole, however, a strategy of investing more savings in bonds usually reduces the volatility of the worker's replacement rate. The offsetting disadvantage is that it increases the contribution rate needed to achieve a given replacement rate. If workers invest all their pension savings in government bonds, the calculations in this chapter imply that they will often obtain lower returns than those available under a fully mature pay-as-you-go pension system.

Note that many workers are likely to prefer a very conservative investment strategy. Evidence on the investment behavior of American workers implies that low-wage workers and workers with limited education tend to allocate their retirement savings to low-risk investment alternatives, including money market funds, bonds and guaranteed income contracts (EBRI 1996; Ameriks and Zeldes 1998). Thus, in comparison with workers earning above-average wages, low-wage workers may experience less fluctuation in the value of their retirement savings, but they would also tend to obtain below-average returns on their contributions and receive below-average pension replacement rates. The investment behavior of low-wage and low-education workers would thus tend to produce even greater proportional inequality in individual-account pensions than is observed in career wages.

Another strategy workers can follow to reduce the uncertainty of their individual-account pensions is 'phased annuitization'. They can convert their retirement savings into annuities over several years rather than at a single point in time, as assumed in the earlier calculations. For example, workers could convert their nest eggs into annuities in regular annual installments beginning several years before they retire. Under one plan, each worker would purchase five annuities rather than only one. The annuities would differ in size depending on stock and bond prices and interest rates at the moment of conversion. Since the conversion occurs in five successive years rather than only once, workers would not convert all their retirement savings into an annuity at a time when asset prices and interest rates make it particularly disadvantageous to do so.

Figure 6.3 shows replacement rates at age 62 under this annuitization strategy for US workers reaching retirement age between 1912 and 2003.



Notes: Assumed contribution rate is 7% of wages. All retirement savings are invested in US equities. Simple annuity is purchased at age 62; phased annuity is purchased in annual installments at ages 58–62.

Figure 6.3 Individual retirement account replacement rates for US workers with alternative annuity strategies, 1912–2003

For purposes of comparison, I also show the replacement rates workers obtain when they convert their retirement savings to an annuity on their 62nd birthday. Both sets of computations assume that 100 percent of pension contributions are invested in stocks.[6] The strategy of phased annuitization yields a distribution of replacement rates that has somewhat less variability, but the strategy also yields a lower average replacement rate.

The standard deviation of replacement rates is 34 percent if the entire annuity conversion takes place at age 62, but it falls to 29 percent when annuitization is phased over five years. The average replacement rate also drops 5 percentage points, however, falling from 71 percent to 66 percent when workers adopt the phased annuitization strategy.

This decline in average replacement rates is hard to avoid. As noted earlier, when workers purchase an annuity they are exchanging stock market investments for a bond market return. By converting part of his/her pension accumulation to an annuity earlier than would be the case if a single annuity were purchased at retirement, the worker who follows a phased annuity strategy is exchanging stock returns for returns on a mixed portfolio of stocks and bonds. This reduces both the variance and the expected return of his/her retirement savings. For the period from

Table 6.3 Distribution of retirement ages under individual retirement accounts with alternative allocations to stocks and bonds

	France	Germany	Japan	UK	USA
100% equities/0% bonds					
Retirement age (% frequency distribution)					
Age 62	50	50	50	50	50
Age 63	11	11	1	14	22
Age 64	11	4	1	8	8
Age 65	7	4	4	3	5
Ages 66–67	13	11	12	16	12
Ages 68–69	5	11	9	8	3
Ages 70–71	4	8	7	1	0
Age 72 or older	0	3	16	0	0
Average retirement age (in years)	63.8	64.4	65.5	63.7	63.2
60% equities/40% bonds					
Retirement age (% frequency distribution)					
Age 62	50	50	50	50	50
Age 63	3	18	5	14	16
Age 64	4	13	3	13	13
Age 65	4	3	5	8	4
Ages 66–67	14	4	14	9	17
Ages 68–69	11	4	4	5	0
Ages 70–71	14	7	4	0	0
Age 72 or older	0	1	14	0	0
Average retirement age (in years)	64.8	63.7	65.1	63.4	63.3

Source: Authors' tabulations as explained in text.

1912–2003, Figure 6.3 shows that the phased annuitization strategy would have done little to improve the worst replacement rates, which are very similar under the two annuitization strategies. Instead, it significantly reduced replacement rates for workers who retired near the end of stock market booms. If the goal of a conservative annuitization strategy is to help workers avoid low replacement rates, the strategy of phased annuitization would have performed poorly for US workers over the twentieth century.

Workers who reach age 62 and have accumulated too few assets to obtain a decent pension can delay their retirement. Suppose workers delay their retirement until they can buy an annuity that replaces at least 70 percent of their peak earnings. How long must workers delay their retirement to achieve this goal? Table 6.3 contains statistics on the distribution of retirement ages under two different retirement saving strategies. As in Table 6.2, the contribution rate for each country has been selected so that half of all workers can accumulate enough savings to attain the target replacement rate by age 62. Consequently, half the workers retire at age 62. Workers who delay their retirement after 62 make pension contributions for one or more additional years (though annual contributions decline slightly, since real earnings are assumed to fall 1 percent a year after age 62). In addition, the price of an annuity would typically decline, since the cost of annuities falls in line with a worker's remaining life expectancy. For US workers the overwhelming majority of retirements can occur by age 65 no matter which strategy is followed with regard to investment of retirement savings. If savings are invested solely in stocks, only 15 percent of retirements must be delayed until after age 65; if a 60 percent/40 percent portfolio of stocks and bonds is maintained, just 17 percent of retirements must be postponed until after 65. Retirements in the United Kingdom are approximately as predictable as they are in the United States.

The variability of investment returns in Japan makes late retirement far more likely in that country. About one Japanese worker in seven fails to accumulate enough retirement savings by age 72 to achieve the target replacement rate, and this is true whether workers hold their savings in Japanese stocks or in a portfolio of 60 percent stocks and 40 percent bonds. Note that a worker who delays his/her retirement to age 72 must supply 25 percent more lifetime labor than a worker who retires at 62. In France and Germany the variability of retirement ages depends on the mix of stocks and bonds in the worker's portfolio. If German workers invest exclusively in stocks they face a greater risk of retirement delays than if they invest in a mixed portfolio of stocks and bonds. In contrast, French workers who select a mixed portfolio face an elevated risk of delayed retirement. In comparison with workers in the United Kingdom and the United States, French, German and Japanese workers face some risk that their retirement will be significantly delayed to achieve a 70 percent replacement rate, and this is true regardless of the investment strategy selected.

7. CONCLUSIONS

Retirement benefits under a private, individual-account system are financed solely with assets held in individual workers' investment accounts. The initial real value of a pension is determined by the current market value of assets held in the account as well as the interest rate at the time of conversion to an annuity. Although proponents of individual accounts are confident that workers can purchase safe assets that will yield high rates of return, the experience of the leading industrial countries suggests that neither the value of financial assets nor their real return is assured. Workers in the same country who follow an identical investment strategy but who retire a few years apart can receive pensions that are startlingly unequal. Investment strategies that produce the highest expected returns and biggest pensions are also strategies that yield the widest swings in pension entitlement. The investment strategy that offers the most predictable pension can produce an expected rate of return that may be lower for many workers than the return available in a pay-as-you-go retirement system.

Even though workers on average can obtain good pensions under an individual-account system, such a system generates wide variability in outcomes, even for workers who follow an identical investment strategy. Assuming that workers deposit 7 percent of their annual pay into a retirement account that is invested in a conservative portfolio of 50 percent Japanese stocks and 50 percent Japanese government bonds, recent experience suggests their initial pensions could range from a low of 24 percent of their peak career earnings up to more than 103 percent of peak earnings. These are the actual replacement rates 62-year-old Japanese workers would have obtained in 1990 and 2003 if their careers and contribution patterns matched the assumptions of this chapter. If an individual-account system were established, the variability of actual outcomes is likely to be even wider than this, because risk-averse workers would choose safer investment portfolios while risk-loving workers would choose portfolios containing a bigger allocation to stocks. The evidence in Figure 6.2 and Tables 6.1–3 provides one explanation for the wider acceptance of funded individual-account pensions in the United Kingdom and the United States compared with France, Germany and Japan. Financial market experience in the English-speaking countries offers workers greater assurance that individual-account pensions will be able to provide initial retirement benefits that are more affordable and in a narrower range than annuities that would be available in the other countries. Nonetheless, workers and policymakers may wonder whether the range of benefit uncertainty is small enough so that individual accounts could provide the foundation for old-age income security, even in Britain and the United States.

Workers who are offered plain-vanilla investment options can follow a couple of strategies to reduce the uncertainty of individual-account pensions. They can gradually convert their retirement savings into annuity income by purchasing annuities at several points in time, reducing the risk that annuities will be purchased at particularly unfavorable prices. They can invest a portion of their retirement savings in bonds or bills, diversifying their investment portfolio. These strategies reduce the volatility of the worker's replacement rate, but they also reduce the expected value of the annuity. If workers invest all their pension savings in lower-risk bills or government bonds, they may easily obtain lower returns than those available under a traditional pay-as-you-go pension system. Perhaps surprisingly, even if workers invest in a portfolio of 60 percent stocks and 40 percent bonds, there is a real possibility in some countries that they will obtain a worse rate of return than the one in a pay-as-you-go system. Finally, risk-averse workers can purchase inflation-protected government bonds timed to mature around the expected age at retirement if such investment products are available.

The introduction of inflation-protected bonds should make it possible for financial intermediaries to offer investment products that assure retirement savers a low minimum real return while allowing savers to participate in some of the upside gains that are possible from equity investments (Bodie and Crane 1999). So far, however, these investment products are not widely available to retirement savers. Some protections are available. Indexed annuities are offered in the United Kingdom and a few other industrial countries. Contributors to the new German individual-account system can purchase investment products carrying a money-back guarantee. This guarantee assures investors of a zero percent nominal return rather than a minimum real return, however.

Financial market turbulence in the past three years (and over the past 13 years in Japan) has reminded savers of the fact that even conservative saving plans offer little assurance that workers will reach old age with enough assets to pay for a comfortable retirement. To some degree the risks can be reduced through improved financial education and the development of new investment products. Although proponents of individual accounts are confident that workers can purchase safe assets that will yield good rates of return, the experience of the leading industrial countries suggests that neither the value of financial assets nor their real return is assured. Workers in the same country who follow an identical investment strategy but who retire a few years apart can receive pensions that are startlingly unequal, a result that most public pension systems are designed to avoid.

To assure income security in old age it makes sense to preserve some components of existing public systems, especially the guarantee of indexed pensions to replace some minimum percentage of lifetime wages. The plans proposed by the US social security reform commission included this kind of guarantee. However, for average-wage workers who choose to participate in the commission's plans, the guaranteed portion of the pension will probably provide less than half of the replacement rate traditionally offered under the social security program. The remainder of retirement income will be derived from annuities or phased withdrawals from new individual retirement accounts (Bosworth and Burtless 2002). Whether this arrangement can provide adequate income security for retired workers is a central question that divides American proponents and opponents of individual-account pensions.

NOTES

* An earlier and shorter version of this chapter was presented at the American Economic Association meetings in Washington, DC, 3–5, 2003, and published in the *American Economic Review, Papers and Proceedings* (May 2003). I am grateful to Claudia Sahm and Benjamin Keys for research help in preparing the chapter. Financial support for this research was received under a grant from Japan's Economic and Social Research Institute (ESRI) to Brookings. The views are solely those of the author and should not be ascribed to ESRI or the Brookings Institution.

1. In the United States, approximately 90 percent of current social security contributions are used immediately to pay for benefits to retired pensioners and their dependents. The contributions needed to finance these benefits must be collected whether the public retirement system is maintained or is replaced by a new system of individual accounts. It is thus incorrect to treat as equivalent the contribution rate to social security and to an individual retirement account. See Geanokoplos et al. (1998).
2. Data on consumer price inflation and on total nominal returns for stock and bond investments were obtained from Global Financial Data in March 2003 (www.globalfindata.com). Bond returns are measured for investments in government bonds with a remaining maturity of at least seven years. Global Financial Data supplies financial information to financial planners, pension funds and investment companies. Where possible, I checked the data against alternative estimates of inflation and stock and bond returns. The information for the United States corresponds very closely to inflation and return data independently derived in a previous paper. See Burtless (2003a).
3. In essence, observations are created for the years 2003–41 based on observed returns for 1927–66. Each annual observation of market returns between 1927 and 2002 is thus used exactly 40 times, once to reflect returns in the first year of a worker's career, once to reflect returns in the second year, and so on up through the last year of a 40-year career. An alternative approach is to predict pensions using Monte Carlo simulation, but this would require specification of the full-time series correlation structure of stock and bond returns, a task that is beyond the scope of this chapter.
4. The results differ slightly from estimates reported in Burtless (2003b). The principal reason is that return data covering the 1927–2002 period rather than 1927–2001 are used

to predict retirement pensions. Both stock and bond returns differed substantially from their historical average values in 2002. In addition, the historical data on government bond yields were modified as a result of small revisions in the Global Financial Data files.

5. Returns are calculated at age 62. Estimated returns would be smaller if they were calculated at the end of a worker's life, assuming that workers convert their retirement savings into an annuity. The expected return from purchasing a level annuity is the bond return, which is substantially lower than the expected equity return (see Table 6.1).
6. Workers who purchase a simple annuity continue to hold all of their retirement savings as US equity investments until they retire on their 62nd birthday. Workers purchasing phased annuities purchase single-life annuities on five successive birthdays starting at age 58. At age 58 they use one-fifth of their savings to purchase an annuity; at age 59 they convert one-quarter of their savings (including the annuity payment from their first annuity; at age 60 they convert one-third of their savings (including the annuity payments from annuities purchased earlier); at age 61 they convert one-half of their savings (including payments from earlier annuities); and at age 62 they convert all remaining savings into a single-life annuity. The results are of course sensitive to the ratios of savings converted on successive birthdays, but strategies that involve larger conversions at earlier ages usually produce lower total pensions at age 62.

REFERENCES

Ameriks, John and Stephen P. Zeldes (1998), 'Empirical evidence on portfolio choice: an analysis of longitudinal data from TIAA-CREF', mimeo, New York: Columbia University Graduate School of Business.

Board of Trustees, Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds (OASDI) (2001), *2001 Annual Report of the Federal Old-Age and Survivors Insurance and Disability Insurance Trust Funds*, Washington, DC: Social Security Administration.

Bodie, Zvi and Dwight B. Crane (1999), 'The design and production of new retirement savings products', *The Journal of Portfolio Management*, **25** (2), 77–82.

Bosworth, Barry P. and Gary Burtless (1998), 'Population aging and economic performance', in Barry P. Bosworth and Gary Burtless (eds), *Aging Societies: The Global Dimension*, Washington, DC: Brookings Institution, pp. 1–32.

Bosworth, Barry P. and Gary Burtless (2002), 'Economic and distributional effects of the proposals of President Bush's Social Security Commission', Boston College Center for Retirement Research conference paper, Chestnut Hill, MA: Boston College Center for Retirement Research.

Brown, Jeffrey R., Olivia S. Mitchell and James M. Poterba (2002), 'The role of real annuities and indexed bonds in an individual accounts retirement program', in J.R. Brown, O.S. Mitchell, J.M. Poterba and M.J. Warshawsky (eds), *The Role of Annuity Markets in Financing Retirement*, Cambridge, MA: MIT Press, pp. 107–52.

Burtless, Gary (2003a), 'Social security privatization and financial market risk: lessons from U.S. financial history', T. Ihori and T. Tachibanaki (eds), *Social Security Reform in Advanced Countries*, London and New York: Routledge, pp. 52–80.

Burtless, Gary (2003b), 'What do we know about the risk of individual account pensions? Evidence from industrial countries', *American Economic Review, Papers and Proceedings*, **93** (2), 354–9.

Dimson, Elroy, Paul Marsh and Mike Staunton (2002), *Triumph of the Optimists:*

101 Years of Global Investment Returns, Princeton, NJ and Oxford: Princeton University Press.
Employee Benefit Research Institute (EBRI) (1996) '*Worker investment decisions: an analysis of large 401(k) plan data*', EBRI Issue Brief No. 176, Washington, DC: EBRI.
Employee Benefit Research Institute (EBRI) (2002), *An evolving pension system: trends in defined benefit and defined contribution plans*', EBRI Issue Brief No. 249, Washington, DC: EBRI).
Geanakopolos, John, Olivia Mitchell and Stephen P. Zeldes (1998), 'Would a privatized social security system really pay a higher rate of return?', in R.D. Arnold, M.J. Graetz and A.H. Munnell (eds), *Framing the Social Security Debate: Values, Politics, and Economics*, Washington, DC: National Academy of Social Insurance/Brookings Institution.
President's Commission to Strengthen Social Security (2002), *Strengthening Social Security and Creating Personal Wealth for All Americans*, Washington, DC: President's Commission to Strengthen Social Security.
US Census Bureau (1996), *Money Income in the United States: 1995*, Series P60–193, Washington, DC: US Government Printing Office.
World Bank (1994), *Averting the Old Age Crisis*, Oxford and New York: Oxford University Press.

7. Pension reforms, tax incentives and saving in Italy

Massimo Baldini, Paolo Bosi, Maria Cecilia Guerra, Carlo Mazzaferro and Paolo Onofri

1. INTRODUCTION

During the 1990s, the Italian public pension system underwent two important reforms, which radically changed its nature. These reforms did not change the general pay-as-you-go framework, but they did change the rules for determination of the benefits. First, there was a shift from earnings-related pensions to a notional defined contribution system. Second, a new indexation rule of pension benefits was introduced: since 1993, benefits have no longer been indexed to changes in the nominal wage rate; they are indexed only to the inflation rate. Meanwhile, the whole decade was characterised by repeated, even if not very successful, efforts to support the start of a private funded pillar (either occupational, or individual) of the overall pension system. In this chapter this experience is evaluated under different profiles.

In Section 2, we study the effects of the pension reforms on the household saving rate and private wealth accumulation. The study is carried out using the Bank of Italy's Survey of Household Income and Wealth. After an analysis of the determinants of the age profile of the household's saving rate and the pattern of financial wealth accumulation, we evaluate the possible effects of the reforms first on the saving rate, with a difference-in-differences approach, and then on the profile of private non-pension wealth, computing the substitution rate between private and net social security wealth. The main conclusions are that, apart from a short-run impact, the saving rate was not significantly affected by the pension reforms, and that the compensating increase in private wealth is about half the change in net social security wealth. This compensating increase is, however, concentrated among the middle and older age groups, while the young do not seem to react significantly to the reforms. On the contrary, there appears to be a negative cohort effect on the saving propensity and on the private wealth accumulation by the young cohorts, which, if projected over the next few decades, would imply a reduction in the aggregate saving rate by four points in excess of the reduction of the equilibrium saving rate.

In principle, tax incentives for private pension systems could help to encourage an increase in voluntary saving. However, as shown in Section 3, it is difficult to find sound theoretical and empirical support for the rationale of these incentives, from the point of view of efficiency and equity. Contrary to the idea that pension funds in Italy do not take off owing to a lack of tax incentives, we show that they are strong enough, even in an international comparison with those in the United Kingdom or the United States. The development of pension funds is hampered by other institutional features of the Italian pension system, and by the difficulties of finding a reasonable settlement that is in the interests of both employers and employees. First of all, the replacement rate of the pay-as-you-go system is still quite high. Second, the Italian worker is provided with a freely disposable capital when he/she quits the firm, he/she is working with. The severance payment fund, *Trattamento di Fine Rapporto* (TFR) (a sort of insurance against unemployment), is actually a compulsory system of precautionary saving, since advances of accumulated savings can easily be withdrawn when important household financial decisions have to be taken. As a consequence, the still high replacement rate, and the precautionary saving function of the TFR discouraged Italian workers from moving significantly from TFR to pension funds, even if the return earned on the severance fund is relatively low. At the same time, for employers, the abandonment of TFR would represent the loss of a secure and relatively cheap source of finance. The main conclusions are that, contrary to a general consensus, in the case of Italy, tax incentives are only one of the instruments that provide protection against the risks associated with old age through a funded pillar, and are not always the most appropriate.

2. REFORM OF THE ITALIAN PENSION SYSTEM AND ITS EFFECT ON SAVING BEHAVIOUR[1]

The Italian Pension System and Its Reforms

Origins: the old generous system

During the first half of the twentieth century, the Italian pension system was a public funded system, where contributions were divided equally among employers, workers and the state. Badly hit by the very high inflation rate during and immediately after the Second World War, it was changed into an unfunded public pay-as-you-go system (pay-go), which was fully developed between 1957 and 1968. These were years of a high rate of economic and demographic growth, which brought about a very generous pension system.

The main features of the pay-go system are the following:

- Pensions were determined by the earnings-related formula: $P = cLW$, where c is the so-called 'internal return coefficient', L is the number of years of contribution (L could not exceed 40), and W is the reference wage. For private sector employees, c was 0.02, and W the average wage of the last five years of work, expressed in terms of the final year prices. The legal retirement age was 55 for women, and 60 for men.
- 'Seniority pensions' were introduced, allowing early retirement for private employees, once they had contributed to the pension system for 35 years, regardless of age. The benefit was computed exactly in the same way as indicated above, with no account being taken of the difference in life expectancy at the different ages of the early retirees.
- Public employees held a privileged position. For several categories of civil servants, the coefficient c was higher than 0.02. For all public employees, the reference wage was the wage of the final year of work. Every civil servant had the right to retire when he/she reached 20 years of 'seniority'; (that is, contribution years); in the case of a woman with two children the minimum seniority for early retirement was 14 years, six months and one day. Again, early retirement pensions were not actuarially fair.
- Between 1957 and 1965, the public unfunded pension system was extended to all the self-employed. Artisans, shop keepers, farmers and professionals had to contribute to their new pay-go funds, but at the beginning they were allowed to receive the benefit as soon as they reached the age of 65, on condition that they had contributed for at least a year. The original pension formula related the benefit to the amount of contribution. In 1990, a new rule was enacted for the self-employed, which granted pension benefits proportional to the average earnings over the last ten years of work with an accrual rate of 2 per cent. Since the level of contributions was not modified accordingly (12 per cent of gross income for the self-employed instead of 27.4 per cent of the gross salary for employees, at that time), this new rule led to a huge growth in future pension net liabilities. Currently, a large share of retired self-employed people still receive a benefit which is subsidied by the state in order to reach the national minimum level of benefit. One should not forget that self-employment represents 30 per cent of overall employment in Italy.
- From 1971 to 1992, pensions were indexed both to price increases and to the average real wage growth.

- The system allowed a uniform substitution rate, but an internal return on contributions variable in the opposite direction from the degree of seniority, the rate of wage growth during the working life, and the age of retirement. The system developed in those years produced a rapid increase in pension expenditure until the first half of the 1990s. The increase in expenditure was also the result of a large number of 'inability' pensions, paid out in lieu of unemployment insurance for workers of a mature age. Figure 7.1 describes the historical dynamics of the expenditure in the past 40 years. The upper line in Figure 7.2, below, shows how expenditure would have grown, in the next 50 years, without any of the reforms enacted during the last decade.

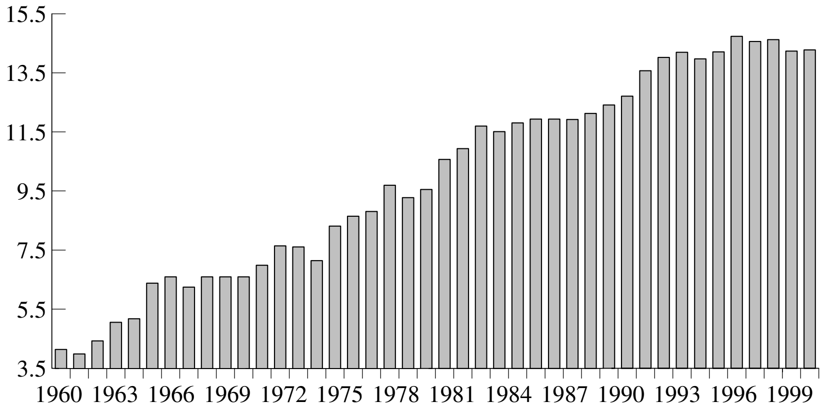


Figure 7.1 Pension expenditure in % of GDP, 1960–2001

First step of the reform: 1992

In 1984 the rules for inability pensions were restricted to the physical inability to work rather than the more general inability to produce income. The long process of reducing the number of inability pensions, which began in 1984 and is still going on, is a component of the slowdown in the growth of pension expenditure. However, it was only in September 1992 that, under the pressure of a deep currency and financial crisis, an effective process of reform started. The first Amato government introduced some important new principles into the system.

- The determination of benefits should adhere more closely to insurance principles through a stronger relationship with the amount of contributions paid during the working life. The reference wage was no longer to be the wage average of the last five years of work, but the average of the last ten years, to be extended, in the future, to the whole working life. The formula for the computation of the first pension benefit for an individual who entered the labour market after 1992 can be expressed as follows:

$$p = 0.02\ LW_0 \left\{ \sum_{i=1}^{L} \frac{(1+w)^{i-1}[1+0{\cdot}01(L)]}{L} \right\}$$

where W_0 is the entry wage; w is the average growth of workers' wages and L is the seniority at retirement.

- Current and future generations can no longer afford to share technical progress with retirees; hence the benefits are no longer indexed to real wages. Moreover, the portion of pension more than three times the official minimum benefit level is no longer fully adjusted to consumer price changes.
- The privileges for civil servants were removed: a very slow process of convergence towards the system for private employees was enacted. The seniority required for a civil servant to retire early was increased.
- The qualifying age for receiving the old-age pension was to be progressively increased to reach 60 years for women, and 65 for men (this target was achieved in 2000). Moreover, the minimum number of years of contribution required in order to receive the benefit was increased from 15 to 20 years.
- Occupational and individual funds should complement the pay-go system, in order to cope with demographic transition. The legislation required was enacted, but the fiscal incentives were not sufficient to enable the funds to be set up.

The measure that made the strongest impact was the abolition of the indexation of benefits to real wage growth. As a whole, our working assumption is that these measures helped workers to realise that the generosity on which they had built their life-cycle wealth expectations was no longer sustainable.

Second step: 1995–97; the notional defined-contribution system

Three years later, in 1995, the Dini government enacted a radical reform of the system of computation of the benefit, based on true insurance principles, applied within a pay-go system. In 1997, the Prodi government further integrated the reform. The new notional defined-contribution (DC) system is a pay-go system that mimics capitalisation. The amount of contributions to the pay-go system during the whole working life will be capitalised at the rate of growth of nominal GDP (actually, a five-year moving average of GDP growth). These capitalised contributions are the basis for the computation of the benefit by means of a discount rate that is a proxy of the expected long-term rate of growth of real GDP (1.5 per cent, in real terms because the benefit is indexed to prices).

Workers can choose to retire at 57; at 65, retirement is compulsory. The computation of the benefit takes into account the average life expectancy (men and women) at the age the worker chooses to retire. Under the new system, the internal rate of return will be the same for every worker, independently of the age of retirement, of the life time wage profile, or of the seniority at retirement. The pension age for men and women will be the same.

Under the new system, the benefit is determined by the following simplified formula:

$$p = a_j W_0 \left[\sum_{t=0}^{T-1} (1+y_n)^{T-t-1} (1+g)^t \right] \gamma(y_r, \omega, s); \text{ with } \frac{\partial \gamma}{\partial y_r} = > 0; \frac{\partial \gamma}{\partial \omega}, \frac{\partial \gamma}{\partial s}, < 0,$$

where a_j is the contribution rate equal to 33 and 20 per cent respectively, for employee and a self employed person, W_0 is the entry wage of an employee of T years of seniority, y_n is the average rate of growth of nominal GDP during the last T years, g is the average rate of growth of the nominal wage during the same T years, and $\gamma(.)$ is the value of a coefficient, which depends on the constant discount rate $y_r = 0.015$ (the expected long-run real growth rate of GDP), on the weighted average of the life expectancy of men and women (ω), and on the probability of having a surviving spouse, combined with his/her life expectancy, s.

The notional DC system introduces actuarial fairness within a pay-go system; the degree of redistribution and solidarity enacted through the pension system is reduced, though not eliminated. The system reduces distortion in labour market decisions, especially with regard to the choice of retirement age. Moreover, it reduces, in a sense, the perverse solidarity of the current system, where workers with a less dynamic career and lower wages finance the benefits of workers with higher and more dynamic wages. A veritable jungle of privileges will be removed: the more favourable treatment both of civil servants (with special reference to the military) and the self-employed compared with private employees will disappear. Under the new system, true solidarity will be mainly a question of general taxation.

If there is no relevant long-run change in income distribution between profits and wages, the new system will tend to balance revenues and payments in a structural way. Given the anticipated real growth of GDP used as a discount rate to compute the benefit, the higher the actual rate of growth, the higher the revenues and the benefits. To fulfil the equilibrium conditions of the system, the coefficients used to compute the benefit will be changed every ten years to take into account changes in life expectancy.

If life expectancy increases, the substitution rate between wage and benefit will be reduced, but the worker can choose to work longer to compensate for such a reduction.

Parametric reforms in 1992 and 1995 dampened expectations for the future level of pension benefits and the change in the indexation mechanism reduced the dynamic of pension benefit after retirement. The first effect can be seen by comparing the replacement rate, defined as the ratio between the first gross pension benefit and the last gross wage. In Table 7.1 we report the replacement rate at different ages of retirement for a representative individual whose seniority at retirement is equal to 37 years. Note

Table 7.1 Replacement rate between pension benefit and last wage at different retirement ages (%)*

	Earnings-related scheme before 1992		Earnings-related scheme after 1992 (Amato reform)		Contribution-related scheme after 1995 (Dini reform)	
Retirement age	E	SE	E	SE	E	SE
58	71.2	67.8	61.9	61.9	54.4	33.0
60	71.2	67.8	61.9	61.9	57.8	35.0
62	71.2	67.8	61.9	61.9	61.7	37.4
65	71.2	67.8	61.9	61.9	68.7	41.6

Notes: *Percentage of the first pension benefit with respect to the last wage of a representative individual who enters the labour market in the year reported in the column and retires with 37 years' seniority. E = employed; SE = self-employed.

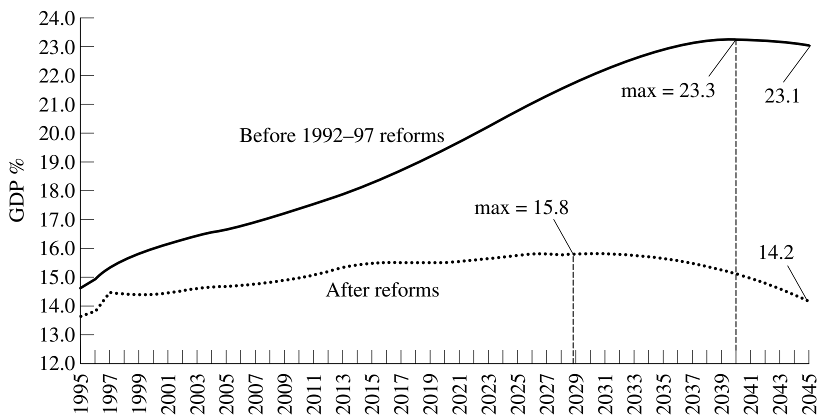
that both the 1992 and the 1995 reforms reduced the replacement rate. With the notional DC formula the replacement rate increases with the retirement age as the formula considers explicitly the average life expectation at retirement. The parameters of the DC scheme are fixed so as to reach the same replacement rate as that of 1992 for a 62-year-old employee. The reduction of the replacement rate is stronger for the self-employed person after 1995 because of the lower level of the contribution rate used to computate the pension benefit, which for this category is fixed at 20 per cent.

Transition period

The new system will not be fully operational for some 25–30 years; therefore the reform has addressed the many details of the long transition process. It has created three different cohorts: young employees, employees with a moderate length of contribution, and mature employees who have a long contribution record. The new system has applied fully to new entrants since the beginning of 1996. For employees who at the end of 1995 had contributed for not more than 17 years, the new system applies only for the remaining part of their working life. Their pension benefit will comprise two parts: the first determined according to the old computation rules, and the second according to the new rules. Finally, for employees who had contributed for 18 years or more in 1995, the rules for the computation of the pension benefit remain the same as pre-1995.

Early retirement benefits will be completely abolished once the new system is fully operational. In the meantime, for workers already at work in 1995, early retirement is progressively restricted to those who are 57 years old and have contributed for 35 years; these conditions will apply in 2004, when the pension rules for both civil servants and private employees will be exactly the same. In other words, early retirement benefit, computed without actuarial fairness, is still allowed for employees who currently have moderate and long contribution records, when they meet the above restricted conditions.

Figure 7.2 describes the behaviour of pension expenditure expected for the next 50 years before and after the reforms of 1992–97. During the first two decades of the century, the reduction in the expenditure dynamics is mainly due to the abolition of the indexation to real wage growth (the strongest effect), to the lengthening of the working life, and to the restrictions in early retirement pensions. When the notional contribution system is fully operational, around 2030, the ratio of the average pension to the average wage will decline (see Table 7.2), bringing back the ratio of pension expenditure to GDP to the current level.



Source: Ministry of the Treasury (1998).

Figure 7.2 Pension expenditures in terms of GDP

Table 7.2 Number of benefits, average benefit, output per employee and number of employees (ratios)

	2000	2020	2030	2050
Average benefit/output per employee	0.158	0.16	0.142	0.11
No. of benefits*/employees	0.900	0.97	1.120	1.22

Note: *The number of pensions is currently 1.2/1.3 times the number of retirees; the change is the result of an increase in the number of retirees.

Source: Italian Ministry of Economy and Finance.

Questions still open

Two main questions are still open in the debate about the reform of the pension system in Italy. The first is how to prevent an already high share of pension expenditure to GDP from further increasing in the next two decades. In other words, how to accelerate the full application of the new regime. This question is not a technical one; it has a deep political value because it implies asking current generations, the same generations that are paying for the reduction of the public debt (that is, for the promises made in the past), to work longer, and to save more in order to pay also for their future. Social and political cohesion will become a very delicate topic in the near future.

The second question is how to support the development of a funded pillar to compensate for the reduction in the expected ratio of public pension to wage, especially for the current young workers. As Section 3 shows, according to the current legislation employers and workers can set up closed (occupational) pension funds. Banks and other financial institutions can establish open funds. Both open and closed pension funds are based on a DC formula. Contributions are tax deductible and derive mainly from a severance pay fund, which is already available. Nonetheless, the development of pension funds has been slow during the 1990s, probably because of the still high replacement ratio offered by the public system to older workers, rather than because of limited tax incentives. Actually, this question too has its social and political counterpart: what is the socially acceptable overall replacement ratio of pension to wages? Once this target is accepted it is easier to decide what weight should be given to the funded part, and how to ask current generations to pay for the cost of opting out of the pay-go system, so that resources to feed the funded pillar can be released. The technical analysis of this process also requires an understanding about how the change in the relative share of private and public social security wealth will influence household saving.

The aim of this research project is to evaluate the impact of the reforms already enacted on the propensity to save both on the general macro level and on the microeconomic behaviour of Italian households. We shall focus on the effects on saving because it is our opinion that every kind of reform that may change the pay-go system and/or stimulate the development of a private funded system will exert a positive effect on the sustainability of future pension claims if it increases the propensity to save, the capital accumulation and the potential growth of the system.

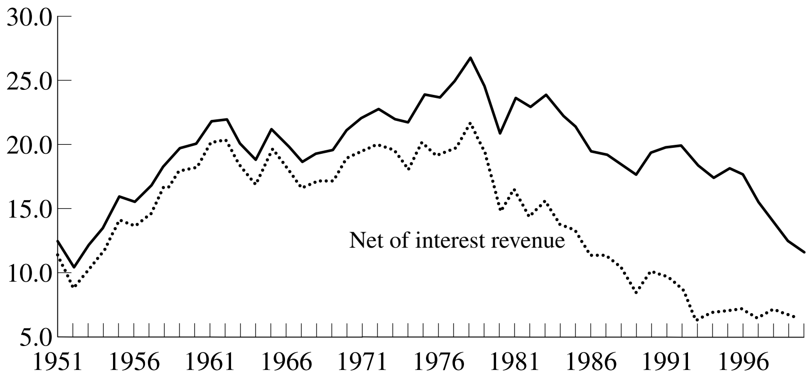
Macroeconomic Effects of the Reforms

The ageing of the population and its decline influence the investment rate, the growth rate of the economy, the demand for financial assets and its composition. Given the rate of capital obsolescence, the *decline* in both population and labour forces releases resources for consumption through a net decumulation of the required capital. *Ageing*, in its turn, implies a stronger conflict of resources available for consumption: higher average propensity to consume, hence a reduction in the net demand for financial assets, with a possible reduction in the average preference for risky assets. A positive effect on saving might be exerted by a transitory implication of ageing: relatively larger and larger mature generations of workers facing smaller and smaller generations of young workers.

It is not easy to state whether the required reduction in saving implied by the declining population, and by the possible change in the rate of technological obsolescence of the capital stock, will be of the same order of magnitude as the reduction in saving implied by ageing. We should not forget that successive generations are not only different in size; they may also differ in their attitude towards saving, as a consequence of the so-called 'cohort effect'. Moreover, as we have already mentioned, different generations and different groups of employees have been affected by the reforms to a different extent. In what follows, we concentrate on the analysis of households' behaviour during the past decade with respect to saving, financial asset accumulation, and its composition in order to try to detect how strong has been the influence of the pension system reforms on their behaviour. In so doing, we hope to be able also to state whether, for the Italian economy, ageing, as such, might imply a reduction in the average propensity to save, and to what extent.

Changes in the propensity to save: the macroeconomic view

From a general macroeconomic point of view, the household propensity to save out of disposable income has been declining since the second half of the 1970s (see Figure 7.3). This long-run trend has been interpreted as the



Note: * See note 2.

*Figure 7.3 Household saving as a percentage of disposable income**

result of two factors. First, a general change in the attitude of households towards saving, fed also by a growing social safety net for household income, has been underway throughout the period. Second, on this long-run effect, several episodes of inflation surprise exerted negative effects on financial wealth, which might have influenced saving decisions. The latter factor has been very much disputed in the past, and to take it into account, household disposable income was computed *à la* Hicks; moreover, possible Ricardian effects in the perception of the 'true' value of household wealth, given the large share of public debt in their portfolio, were explored. The estimations of both the Hicksian and Ricardian effects revealed quite low coefficients to correct the current computation of household disposable income. Identification of those effects was difficult, given the possible effects exerted by changes in disposable income distribution, due to the high degree of concentration of wealth ownership and of the resulting yields.

If we restrict our analysis to the last decade, some of the phenomena that influenced household wealth in the previous decades reversed: inflation almost disappeared, public debt started to decline, the social safety net stopped widening and, finally, private financial wealth earned large capital gains. Nonetheless, the propensity to save continued to decline. Only if we include capital gains in household disposable income does propensity to save increase in the second half of the 1990s (see Zollino 2001). However, if we simply compute household disposable income net of interest revenues, the trend to decline stops after 1993.[2]

Thus, too many other macroeconomic factors have been at work for us to be able to detect the impact of the pension reforms at the aggregate level. On the one hand, the government budget was consolidated: between 1992 and 1997 the general government deficit was reduced from 10.7 per cent of GDP to 2.7 per cent. The currency crisis of summer 1992 was the signal that current and future taxes plus a reduction in public expenditures would reduce the forward-looking part of the life-cycle wealth of Italian households, which might have exerted an expansionary effect on the propensity to save. On the other hand, the supply of financial assets changed dramatically: the liberalisation of the capital markets, the reduction in the supply of government bonds, the bubbles in the equity markets and the privatisation of state-owned firms are other factors that may have influenced the behaviour of saving. We need to explore further the microeconomic household behaviour, and will start with the age profile of saving.[3]

Age profile of the propensity to save

Most empirical results show a very slow decline in the propensity to save with age (Ando et al. 1994). The usual explanation is that, up to now, the old generous pension system has guaranteed a high substitution rate between wage and pension benefit, and the computation of the saving rate does not take into account the decumulation of individual social security wealth as pension annuities are paid (Jappelli and Modigliani 1998). This is usually supplemented by the presence of a bequest motive, precautionary saving, and the uncertainty about the length of life. As a first stage, we shall ignore the effect of individual social security wealth, and use the data from the Bank of Italy's Survey of Household Income and Wealth (SHIW), which refers to six surveys of income and wealth of Italian households, carried out in 1989, 1991, 1993, 1995, 1998 and 2000. On average, each survey covered about 8000 households.[4]

In general terms, the age profile of a behavioural variable depends not only on a pure age effect, but also on possible cohort effects, and on the possible presence of time-period effects. The cohort effect is the result of the fact that different generations behave in systematically different ways at the same age. The time-period effect results in a systematically different behaviour at all current ages at a given moment of time; in other words, it shifts either up or down the age profile, without changing its shape. If a is the age, t is the calendar time, and c the year of birth of a given cohort, $a = t - c$. The collinearity among these three effects does not allow their separate identification. As a consequence, in order to obtain an estimation of the age profile we need to use an identification restriction on the coefficients of both simple cross-section regressions, and panel cross-section regressions. The shape of two of the three effects can be estimated only as conditional on an exogenous shape of the third one. To give an example, moving from a survey at time k to the following one at time $k+1$, if we observe an equal percentage increase in the average value of a given variable for each age and cohort, it might be the result either of a trend effect or a combination of a positive age effect, and of a cohort effect favourable to the younger generations. Here we shall use two different sets of restrictions. In the first one, the variability of the data is assumed to be fully imputable to a combination of age and time-period effects; in the second, we restrict to zero the possible time-period effects, in order to obtain an estimation of the cohort effects.

In order to isolate the age effect, we have also derived a pseudo-panel from the original 47 834 household observations of the pooling of the six surveys. The pseudo-panel describes the time evolution of the average behaviour of homogeneous groups of households. These groups are defined by the year of birth of the 'breadwinner'.[5] Given the limited number of surveys used, we divided the sample into 12 cohorts: the first cohort includes all households whose breadwinner was born between 1910 and 1914, the twelfth cohort those born between 1965 and 1969. The limited temporal extension of the sample produces little overlap between different cohorts over time, which is very important for estimating the cohort effects; choosing cohorts defined on a period of less than five years would, however, reduce the probability of detecting significant systematic differences across successive cohorts. For each cohort we have computed the average value of the economic variables describing their behaviour; the values have been expressed in terms of 1998 prices, using the consumer price index. The saving rate has been defined as the difference between disposable income and consumption, divided by disposable income. Consumption is net of durable services and capital gains.

An estimate of the various effects involved can in principle be obtained, without significant differences, with regressions either on the individual micro data or on the pseudo-panel of their cohort averages. We have followed both options, but the main text describes, for space reasons, only the results obtained from individual data; results obtained from the pseudo-panel are not qualitatively different. Table 7.3 reports the estimation results, from the individual micro data, under the two alternative identification restrictions: first, without cohort effects, and then without time dummies. The omitted dummies refer to the first cohort and to the first available year, 1989. The regression method is ordinary least squares (OLS), with observations weighted by the sample weights provided in the survey.[5] Standard errors have been computed with the White correction for heteroscedasticity. The table shows the coefficients estimated from regressions on the individual micro data, carried out both with and without a set of demographic controls, which may influence the evolution of permanent income and therefore of personal saving over the life cycle.

With regard to time effects, both sets of regressions provide consistent results: there is no definite trend of the saving rate over the available period.

Table 7.3 Regression results for the saving rate: micro data – age and time or cohort effects

	Age and year effects				Age and cohort effects			
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Age	0.0021	0.0003	0.0021	0.0003	0.0017	0.0006	0.0006	0.0006
Age^2	−0.0001	2.25E-05	2.29E-05	2.12E-05	−0.0000686	0.0000616	3.34E-05	0.0000616
Age^3	−2.56E-06	6.85E-07	−8.20E-07	6.50E-07	−3.35E-06	−1.34E-06	9.73E-07	−1.34E-06
Age^4	8.48E-08	3.63E-08	−2.21E-08	3.40E-08	6.27E-08	−4.10E-08	4.67E-08	−4.10E-08
1991	0.0255	0.0047	0.0202	0.0044				
1993	0.0002	0.0052	−0.0067	0.0049				
1995	−0.0454	0.0050	−0.0553	0.0047				
1998	0.0152	0.0055	0.0036	0.0052				
2000	−0.0035	0.0051	−0.0147	0.0049				
Cohort2					−0.0047	0.0119	−0.0084	0.0106
Cohort3					−0.0147	0.0112	−0.0149	0.0104
Cohort4					−0.0240	0.0129	−0.0252	0.0118
Cohort5					−0.0318	0.0143	−0.0377	0.0132
Cohort6					−0.0165	0.0154	−0.0292	0.0143
Cohort7					−0.0188	0.0164	−0.0362	0.0153
Cohort8					−0.0273	0.0172	−0.0491	0.0161
Cohort9					−0.0312	0.0183	−0.0555	0.0172
Cohort10					−0.0360	0.0193	20.0687	0.0182
Cohort11					−0.0447	0.0205	20.0857	0.0194
Cohort12					−0.0529	0.0226	−0.0963	0.0215
Primary			3.15E-02	4.08E-03			3.08E-02	4.15E-03
Secondary			0.0707	0.0043			0.0708	0.0043

Tertiary			0.1221	0.0058			0.1237	0.0058
Centre			−0.0262	0.0038			−0.0261	0.0038
South			−2.64E-02	0.0033			−0.0275	0.00339
Male			−9.84E-04	3.70E-03			−1.89E-04	3.76E-03
Self-empl.			0.0034	0.0047			0.0050	0.0048
N earners			0.0893	0.0019			0.0886	0.0020
Const.	0.2420	0.0038	0.0507	0.0063	0.2650	0.0162		
N obs.	43971		43971		43971		43971	
R^2	0.0160		0.1380		0.0060		0.1310	

In 1995, the saving rate is clearly lower than in adjacent years, and in 1998 it is almost at the same level as at the beginning of the decade. The estimations with age and time dummies show alternatively positive or negative time coefficients (to be interpreted as deviations from the 1989 average), with a slightly declining tendency. The regressions also show that the saving rate is higher in the northern part of Italy, the richest one, and that it is very positively correlated with the education of the household head. The self-employed do not appear to have saving rates different from the rest of the working population. Turning to the cohort effects, they seem to be precisely estimated, both with and without demographic controls, and highlight the presence of a negative effect for the younger generations: at the same age, younger cohorts seem to save less than the older ones. The different impacts on saving behaviour exerted by single cohort and age effects are shown in Figures 7.4–7.6.

Figure 7.4 shows the evolution of the saving rate over the life cycle; the data are organised as ratios of the averages of saving and disposable income for each cohort in each survey.[7] The various lines describe the actual time behaviour of the propensity to save in the interval 1989–2000, for each cohort. The propensity to save is increasing up to the age of 65, and then declines only in very old age. In any case, saving is positive at all ages. Of course, should we consider pension benefits not as income, but as a de-

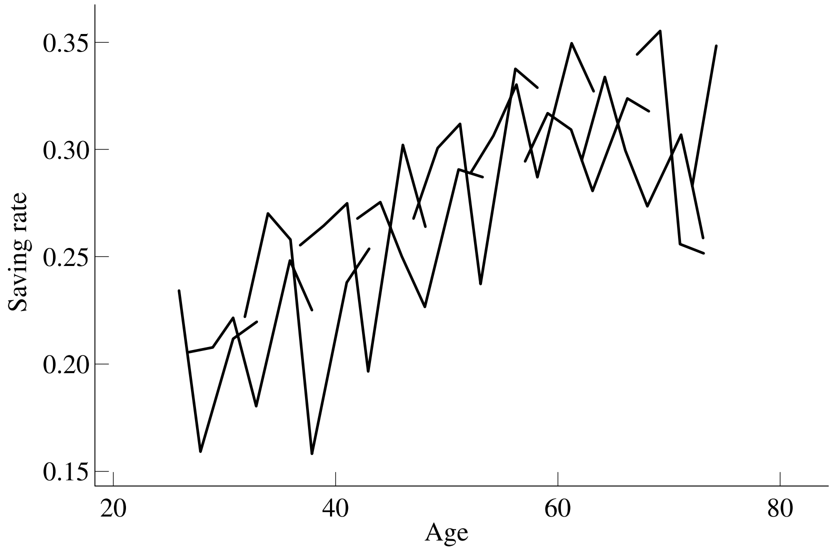


Figure 7.4 Propensity to save: cohort averages

cumulation of the individual social security wealth, and should we compute saving as the change in the total wealth (including social security wealth), the propensity to save of the elderly would be much lower. The 1998 survey (the year of the end of government budget consolidation, and of strong recovery of the equity market) shows a strong increase in the propensity to save for each cohort. In 2000, saving rates of most cohorts remain fairly close to their 1998 levels.

Figure 7.5 Propensity to save: age effect estimated from the regression on age polynomial and time-period dummies

Figure 7.5 describes the age effect resulting from a regression on the polynomial on age and time-period dummies (without the demographic controls). The level of this profile depends on which time-period dummy is dropped from the regressions; the shape of the profile, however, is by construction unique for each period. It confirms the shape observable in Figure 7.4, with a very limited decline at old ages.

Figure 7.6 shows a profile that derives from a regression on the age polynomial and the cohort indicators (again from Table 7.3). The cohort effects seem to be negative for the younger generations. As younger breadwinners enter the sample, they show a propensity to save, at a given age, lower than the one shown by previous generations at the same age. Ideally, if complete panel data sets for all the households observed were available, we would observe parallel age profiles, with those of younger cohorts placed below

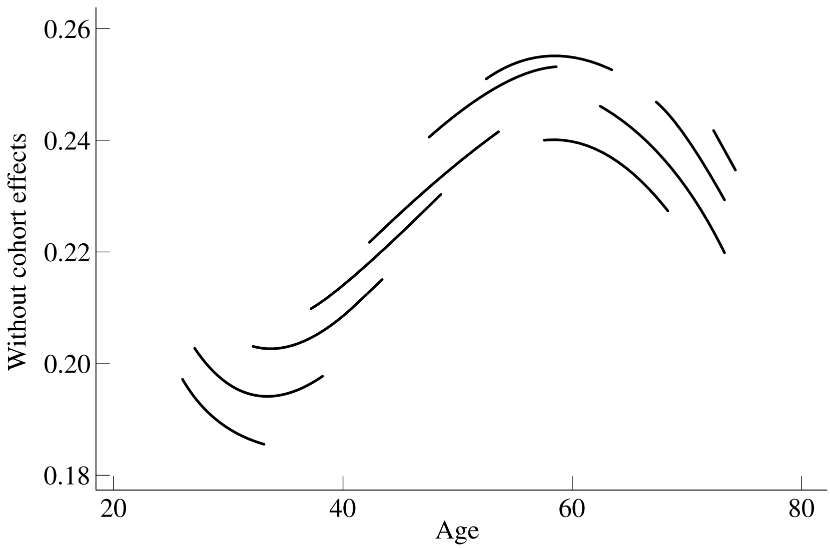


Figure 7.6 Propensity to save: age profile estimated with cohort dummies

those of the older ones; this is of course a strong simplification of reality, but it could be useful if it were actually able to capture a possibly important phenomenon taking place in the last few years, namely a tendency by younger cohorts to reduce their propensity to save. Finally, both identification assumptions (age polynomial with time-period dummies, and age polynomial with cohort dummies) reveal the same age effect.

Together, these results suggest that over the 1990s, in spite of the pension reforms, and at a microeconomic level of analysis, the household saving rate showed a tendency to continue the decline of the previous decade. The decline appears more intense for younger cohorts, though they are the cohorts affected most by the pension reforms of the decade.

Ageing and financial assets accumulation

The other side of the coin of the age distribution of the propensity to save is the age distribution of personal wealth. A section of the SHIW includes detailed information on the household portfolios. Unfortunately, during the 1990s, the surveys changed the definition of financial assets. As a consequence, it is not so easy to build detailed and homogeneous times series of the households' portfolio choices.

For example, in the 1998 Survey, mutual investment funds are considered as a single item of the portfolios independently of the degree of risk their assets may imply: mainly equities, mainly bonds, or a mix of both. This requires working with very wide definitions of the financial aggregates, such as: (i) *AF1*: Deposits with both bank and postal systems; (ii) *AF2*: Italian Treasury bonds with very different maturities; (iii) *AF3*: other financial assets.

AF1 and *AF2* can be considered as risk-free assets, whereas *AF3* includes very different assets with different degrees of risk. Table 7.4 shows the average and median values of the whole stock of financial assets held by households, and the change in their portfolio composition along the time period of the data (1989–2000). As time elapses, *AF3* increases its weight significantly. While we should not forget that government bonds can also be held through mutual investment funds, an increasing tendency to hold more risky assets seems to have emerged during the past decade.

Table 7.4 Composition of household portfolios

	Average of total financial assets (€000 1998 prices) *AF*	Median of total financial assets (€000 1998 prices *AF*	Bank and postal deposits *AF1/AF*	Italian Treasury bonds *AF2/AF*	Other financial assets *AF3/AF*
1989	16.9	3.8	0.61	0.29	0.10
1991	15.4	5.7	0.55	0.31	0.13
1993	18.4	5.7	0.46	0.32	0.23
1995	18.3	5.5	0.44	0.37	0.19
1998	24.2	7.7	0.49	0.12	0.39
2000	26.7	5.7	0.47	0.14	0.39

Table 7.5 shows that during the same period the number of households holding at least one kind of financial asset has increased. Nevertheless, in 2000 about 19 per cent of households do not hold any kind of financial assets. The change that we observe in household preferences seems to be a structural one, including both a growing preference for holding financial assets *vis-à-vis* real estate properties and a stronger tendency towards risky assets.

Also in this case, we wonder whether the general macroeconomic effects allow a precise trend effect on the behaviour of the ratio of financial assets to household disposable income to be detected.

Moving to the analysis of time-period, age and cohort effects on the age profile of asset holding, we should remember that financial survey data include all three effects. The propensities to hold financial assets and risky assets, as pure results of different ages, cannot therefore be isolated from the data of a single time-period survey. The first step is the estimation of

Table 7.5 Percentage of households holding financial assets

	Bank and postal deposits *AF1*	Italian Treasury bonds *AF2*	Other financial assets *AF3*	Total financial assets *AF*
1989	68.9	18.0	6.0	68.9
1991	80.8	23.2	7.5	81.1
1993	82.6	22.4	9.7	82.9
1995	83.4	26.2	10.5	83.8
1998	86.0	11.8	18.1	86.1
2000	80.4	11.8	21.1	81.0

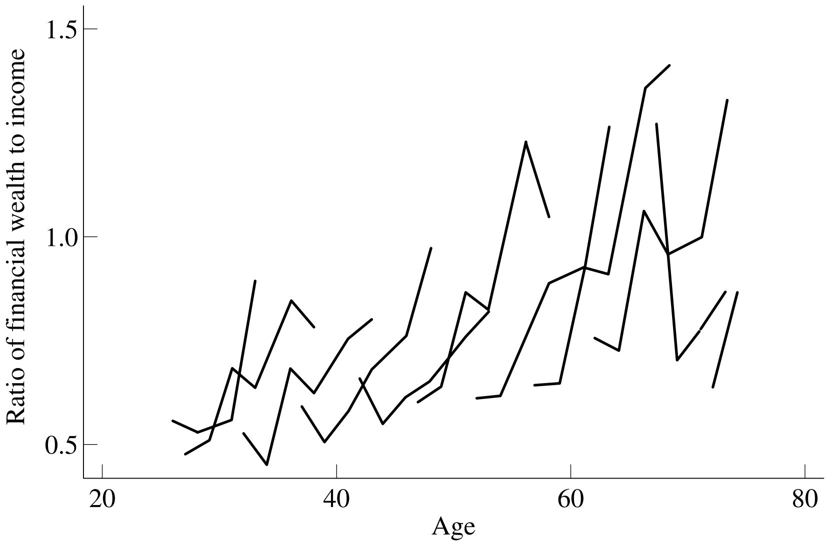


Figure 7.7 Ratio of total financial assets to annual disposable income

the age profile of the ratio of the stock of financial assets (*AF*) to the disposable income (*Y*), *AF*/*Y*. Figure 7.7 shows its behaviour according to age, and along the six surveys taken into account. Each line connecting a set of six points refers to the *AF*/*Y* ratios for a single cohort.

During the ten years of the surveys, all cohorts show a systematic increase in the *AF*/*Y* ratio. In the last part of the period, the ratio increases to a larger extent for almost all generations, and particularly so for those in the second part of their life cycle. As we have already mentioned, during the last and the last but one survey the liberalisation of the capital markets, the large market value of the privatisation of state-owned firms, the huge capital gains on the equity market and the sharp reduction in the inflation rate have increased both the value of financial assets and the propensity to hold them. Moreover, the more mature generations of workers and the generations of retirees have increased their *AF*/*Y* ratio to a larger extent. General macroeconomic phenomena combined with specific age and cohort behaviour seem to have been at work. A very simple inspection of Figure 7.7 suggests that the time-period effect is spread to all cohorts, even if to a different extent. This effect is very strong in 1998 to overcompensate the declines in the previous years.

At this point, the usual problem of separating age, time and cohort effects emerges. One of the most common procedures (see Deaton and Paxson 1994) followed by researchers is to ensure that the coefficients of time-period dummies sum to zero, hence that they are serially uncorrelated random shocks. We have also tried this identification restriction, but from the regression of the *AF*/*Y* ratio on two fourth-degree polynomials – one in the age, and one in the cohort index – and on time-period dummies, whose coefficients are restricted to sum to zero, we have obtained an age profile for *AF*/*Y* always increasing during the life cycle, compensated by large and increasing cohort effects for the younger generations. These results are not shown, since the magnitude of the cohort effects that can be obtained in this way is so large as to induce us to believe that a trend effect covering the whole sample period is dominant in these data, so it would be misleading to impose the absence of such a trend.

We thus try an alternative identification assumption (similar to one of the two used before for the saving rate), used by Paxson (1996) for the estimation of the age profile of the saving rate, and by Guiso and Jappelli (2000) for the estimation of the share of households holding risky assets: if we arbitrarily assume that different generations have no different attitudes towards financial assets, we can ignore possible cohort effects and regress our dependent variable both on age and on unrestricted time-period dummies. The age profile that we get from this assumption is presented in Figure 7.8. The absolute level of the profile comes from excluding the dummy for 1989 from the regression.[8] Moreover, the absolute level of the ratio is quite low, about a third of the amount of financial wealth emerging from the macroeconomic data.

The difference has to be imputed to the typical underestimation in the answers to a survey, when personal wealth level is concerned. Actually, we are interested in the behaviour of the ratio during the individual life cycle, and during the time period, not in its absolute level. We must proceed as if the degree of the under estimation were the same at all ages, even if we understand that it may be a function of the level of wealth, and the latter is not independent of age.

Figure 7.8 Ratio of total financial assets to disposable income: age profile estimated without cohort effects

The a priori exclusion of possible cohort effects is a strong assumption.[9] The attitude towards holding financial assets *vis-à-vis*, for example, real estate, may have changed with successive generations. The alternative solution, corresponding to the second restriction that we have already used in the estimation of the saving rate, used by Venti and Wise (1993), and more recently by Poterba (2000), excludes the time-period effects, through the use of polynomials (or equivalently sets of dummies) both in the age and the cohort indices. Of course, this implies the risk of interpreting general macroeconomic phenomena common to all generations as cohort effects.[10] We have also tried this alternative (very similar to imposing estimated time coefficients but restricted to zero). The age profile that we obtain (from the right-hand side of Table 7.6) is continuously increasing, and shows relevant cohort effects. These are very large, and increasing mainly for the younger generations. The suspicion that age and cohort effects which emerge as very relevant, and with opposite sign are, actually, the result of a general trend in the data cannot be ignored. This is the same conclusion reached by Ameriks and Zeldes (2000), and by Guiso and Jappelli (2000) in very similar contexts. We, however, present results from both alternative regressions, AF/Y on age and time effects, and AF/Y on age and cohort effects, in Table 7.6, which provides the same information as Table 7.3, above, for the saving rate: the estimates of age, cohort and time coefficients for the ratio *AF*/*Y*, over the same sample period, according to these two alternative identification restrictions. The pattern of cohort and period effects is fairly clear: there is an increasing tendency to invest in financial assets, particularly by the younger generations. This is probably due to a time effect deriving from the structural changes mentioned before. Turning again to the individual regressions (Table 7.6), the coefficients of year dummies, where present, are always very significant and of increasing magnitude; cohort coefficients are clearly influenced by the presence of a trend, leading to a rapidly increasing pattern of wealth-holding for younger cohorts. The socio-demographic controls show that more-educated households own a much greater amount of financial assets, and that households living in the North are richer than those living elsewhere. Being self-employed is associated with a higher ratio of financial assets to disposable income. The values of time and cohort coefficients do not change very much with the inclusion of these additional regressors.

The age profile of total financial asset holding is slightly increasing, and it declines only in the very last part of the life cycle. The coefficients of the second column of Table 7.6 (age and year effects in the individual data, without demographic controls) have been used to draw the continuous line in Figure 7.7, above, relative to the pure age effect.

Summing up, the data seem dominated by a structural change in attitude of all generations towards a greater propensity to invest in financial assets; as a result, the ratio of *AF*/*Y* increases up to age 70, before declining, but very slowly.

Ageing and attitudes towards risk

We shall now discuss the age profile of the portfolio composition, focusing on the share of risky assets held by households. We first show the regression results for the various effects; then we draw a figure with the original data and the age effect obtained from the regression on individual data without cohort effects (that is, using the coefficients in the second column of Table 7.7).

Like the ratio *AF*/*Y*, *AF3*/*AF* has also experienced a strong growth throughout the whole decade, especially in the 1995–98 period. As a consequence, we ran both types of regression as for *AF*/*Y*: the first one without cohort effects (Figure 7.9) and the second without the time period effects (Figure 7.10). Also in this case, the latter estimation produces age and cohort profiles that are very large in size and with opposite signs (if we plot them in the same graph against age, the age profile is increasing with age, and the level of cohort coefficients is much higher for the younger generations). The age profile we get from the estimation without cohort effects

Table 7.6 Regression results for the ratio AF/Y*: micro data – age and time or cohort effects*

	Age and year effects				Age and cohort effects			
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Age	0.0076	0.0010	0.0159	0.0010	0.0164	0.0022	0.0228	0.0022
Age^2	−1.38E-04	7.29E-05	−1.24E-04	7.07E-05	−5.52E-05	0.000124	−7.44E-07	0.00012
Age^3	−6.32E-06	2.21E-06	−1.10E-05	2.22E-06	−1.45E-06	2.97E-06	−5.82E-06	2.98E-06
Age^4	8.91E-08	1.03E-07	1.30E-07	9.94E-08	2.95E-08	1.34E-07	4.04E-08	1.29E-07
1991	−0.0127	0.0209	−0.0108	0.0208				
1993	0.0510	0.0208	0.0623	0.0205				
1995	0.0835	0.0216	0.0881	0.0212				
1998	0.1728	0.0245	0.1519	0.0243				
2000	0.0910	0.0232	0.0823	0.0225				
Cohort2					0.1371	0.0507	0.1223	0.0486
Cohort3					0.2335	0.0552	0.2029	0.0534
Cohort4					0.3109	0.0601	0.2878	0.0579
Cohort5					0.3882	0.0702	0.3644	0.0673
Cohort6					0.4176	0.0750	0.3781	0.0723
Cohort7					0.5202	0.0823	0.4636	0.0794
Cohort8					0.5017	0.0819	0.4522	0.0789
Cohort9					0.5521	0.0848	0.4957	0.0817
Cohort10					0.6105	0.0875	0.5271	0.0844
Cohort11					0.7069	0.0887	0.6047	0.0859
Cohort12					0.6948	0.0914	0.5997	0.0886
Primary			0.1416	0.0172			0.1449	0.0173
Secondary			0.2900	0.0186			0.2933	0.0186

Tertiary			0.4806	0.0278			0.4835	0.0278
Centre			−0.2188	0.0146			−0.2180	0.0145
South			−0.3432	0.0140			−0.3416	0.0139
Male			0.1086	0.0149			0.1057	0.0148
Self-empl.			0.1881	0.0192			0.1879	0.0193
N earners			−0.0746	0.0084			−0.0741	0.0085
Const.	0.5388	0.0181	0.5425	0.0307	0.0971	0.0800	0.1477	0.0809
N obs.	46150		46150		46150		46150	
R^2	0.011		0.077		0.010		0.076	

Table 7.7 Regression results for the ratio AF3/AF*: micro data – age and time or cohort effects*

	Age and year effects				Age and cohort effects			
	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.	Coef.	s.e.
Age	−0.0004	0.0002	0.0011	0.0002	0.0105	0.0005	0.0115	0.0005
Age^2	−4.31E-05	1.26E-05	−1.46E-05	1.23E-05	−6.17E-05	2.16E-05	−2.49E-05	2.05E-05
Age^3	−1.29E-07	4.32E-07	−7.37E-07	4.22E-07	−1.35E-06	7.04E-07	−1.88E-06	6.73E-07
Age^4	−1.02E-09	1.67E-08	−1.00E-08	1.60E-08	−9.38E-09	2.26E-08	−2.29E-08	2.14E-08
1991	0.0118	0.0026	0.0109	0.0026				
1993	0.0309	0.0030	0.0318	0.0030				
1995	0.0335	0.0028	0.0324	0.0028				
1998	0.0940	0.0044	0.0882	0.0042				
2000	0.1190	0.0046	0.1148	0.0044				
Cohort2					0.0187	0.0068	0.0153	0.0064
Cohort3					0.0441	0.0074	0.0387	0.0070
Cohort4					0.0845	0.0090	0.0801	0.0085
Cohort5					0.1366	0.0107	0.1313	0.0100
Cohort6					0.1895	0.0121	0.1796	0.0114
Cohort7					0.2557	0.0133	0.2414	0.0125
Cohort8					0.2994	0.0138	0.2849	0.0130
Cohort9					0.3467	0.0148	0.3310	0.0141
Cohort10					0.4075	0.0156	0.3853	0.0148
Cohort11					0.4678	0.0170	0.4401	0.0161
Cohort12					0.5152	0.0194	0.4880	0.0184
Primary			0.0273	0.0032			0.0280	0.0032
Secondary			0.0709	0.0036			0.0723	0.0036

Tertiary			0.1342	0.0065			0.1360	0.0066
Centre			−0.0411	0.0032			−0.0408	0.0032
South			−0.0685	0.0023			−0.0674	0.0023
Male			0.0123	0.0027			0.0119	0.0027
Self-empl.			0.0374	0.0042			0.0377	0.0042
N earners			0.0097	0.0016			0.0093	0.0016
Const.	0.0267	0.0020	−0.0168	0.0044	−0.1953	0.0126	−0.2275	0.0126
N obs.	46394		46394		46394		46394	
R^2	0.055		0.131		0.049		0.126	

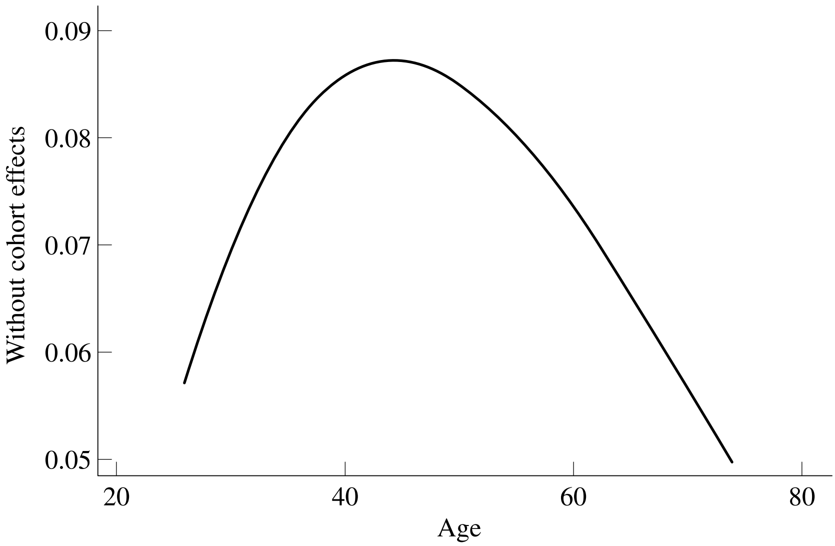


Figure 7.9 *Ratio of risky assets to total financial assets: age profile estimated without cohort effects*

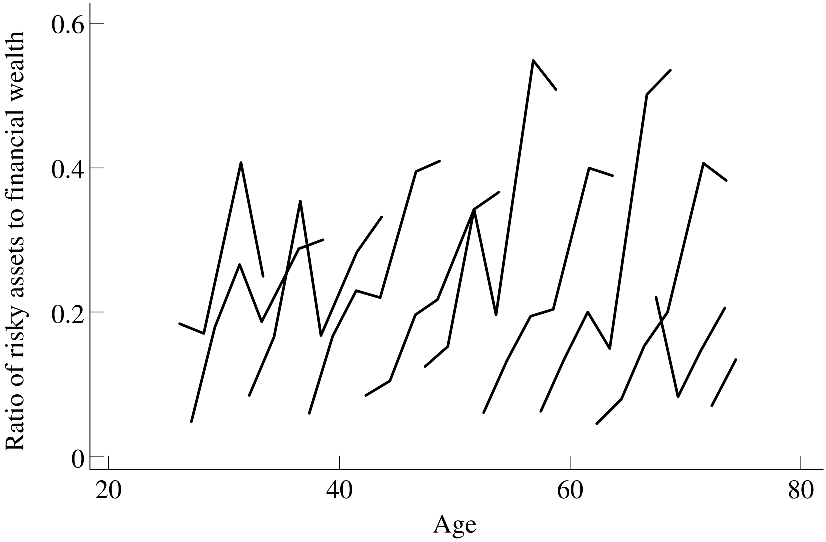


Figure 7.10 *Ratio of risky assets to total financial assets: age profile estimated without time-period effects*

shows first an increase in the share of risky assets, and then a reduction after the age of 50.[11] This is common knowledge for financial promoters, even if both Ameriks and Zeldes (2000) for the US and Guiso and Jappelli (2000) for Italy have shown that the reduction in the share of risky asset holding with age, on average, is the result of two different phenomena: a hump-shaped age profile for the *holding* of risky assets as such, and a more or less stable profile for the *share* of risky assets, conditional on holding them. Also in this case, our opinion is that the age profiles estimated without cohort effects seem more reliable.

Summing up, data appear to be dominated by a trend effect: the holding of risky assets has strongly increased during the period under observation. Cohort effects show that such an attitude towards holding risky assets was stronger for the younger generations. Both these trend and cohort effects might distort the evidence on the attitude towards risk at an older age.

Microeconomic Effects of the Reforms: The Difference-in-differences Approach

To sum up the results we obtained above:

- the aggregate propensity to save has been declining only slightly during the past decade, after two decades of stronger decline;
- the age profile of the propensity to save seems to have been influenced more by cohort effects than by general trend effects (however, when the time-period effect is more precisely detected it seems to have reduced individual propensity to save);
- the age profile of the ratio of financial assets to household disposable income has been subject to relevant trend effects over the last decade; and finally,
- the age profile of the holding of risky assets has also shown a marked shift upwards.

As long as the substitutability between real and financial wealth, on the one hand, and pension wealth, on the other, is different from zero, the picture which emerges from the previous section is incomplete. Here, we examine in more detail the role played by the reforms of the pay-go system of the last decade, focusing on the saving behaviour of very specific groups of households, affected to a different extent by the reforms.

The theory suggests that there should be a linkage between the generosity of the mandatory pension system and the level of voluntary saving: the higher the standard of living that the state guarantees in old age, the lower should be, *ceteris paribus*, the need to save for retirement. A reduction in the social security wealth determined by the reforms should therefore induce households to increase savings.

To this end, it is convenient to define individual saving as the sum of compulsory and voluntary saving, where the former measures the participation in both pay-go and funded pension schemes (Jappelli and Modigliani 1998; Miles 1999; Borsch-Supan 2001). Some authors have, accordingly, argued that in the estimation of the age profile of saving, the contributions to a pay-go scheme should be measured as positive saving and pensions as a rent. Jappelli and Modigliani (1998) and Brugiavini and Padula (2001) use this definition of saving for Italy and find that the sum of the estimated compulsory and voluntary saving produces an age profile much more consistent with the one predicted by the life-cycle hypothesis (LCH). Therefore they argue that the substantial flat profile of voluntary saving estimated from the data should not be considered as a rejection of the LCH.

The degree of substitutability between compulsory contributions to a pay-go system and voluntary saving is also important for explaining the effects of pension policies on private saving. An increasing number of papers have estimated such a degree of substitutability for specific cohorts. Gokhale et al. (1996), for example, find that the generosity of the American social security system explains a large part of the decline in private saving for the generations which had gained more from these policies. Alessie et al. (1999) find that productivity growth and social security can explain most of the differences in wealth holdings among living cohorts in the Netherlands. They find that the cohorts that had social security throughout their life have less than half of the accumulation rate of older cohorts. As we have also seen in Italy, the 1970s witnessed the growth of a very extensive social security system which realised a significant intergenerational transfer of resources in favour of the generations who retired between the beginning of the 1970s and the end of the 1980s (Castellino 1995). The high replacement rate of pension to final wage and the generous indexation of pensions to real wage growth until 1992 have determined a growth of the social security wealth from 40 per cent of GDP in 1980 to 270 per cent of GDP in 1992 (Beltrametti 1997).

According to the standard LCH, the generations who received this transfer of lifetime resources should have decreased the accumulation of voluntary saving and therefore also of private wealth. The available empirical evidence suggests that for Italy as well, the degree of substitutability between these two forms of saving is far from unity: on Italian micro data, Brugiavini (1987) found a substitution rate of only 10 per cent between private and social security wealth, and Jappelli (1995) found values around 16–20 per cent. In this and the following section we shall provide new evidence on this topic in two different steps: first, by estimating the effects of pension reforms on the propensity to save of different socio-demographic groups, using the *difference-in-differences* approach; and second, by computing social security wealth for each individual of the six surveys (1989, 1991, 1993, 1995, 1998, 2000), and estimating the effects of pension reform on private wealth accumulation.

Social security reforms and household saving rate under the difference-in-differences approach

The reforms of the Italian public pension system provide a very interesting opportunity to test the degree of sensitivity of private saving to changes in social security wealth. They did not have a uniform impact on the population, reducing pension wealth very unevenly across ages and social groups. In particular (see Castellino 1995; Beltrametti 1997), the reforms reduced the rights of those who were already retired, at the moment of their introduction, only through the general abolition of real wage indexation, and also guaranteed a long transitory phase to preserve the position of those workers who had a contributory record of at least 15 years of work in 1993. The reforms apply fully only to those who entered the labour market after 1995. Other significant distribution effects cut through various demographic and professional groups. For example, the introduction of a close connection between contributions paid and future pension has strongly negative effects on the future benefits of the self-employed, who are still subject to lower contribution rates than the employees. On the other hand, public sector workers have been penalised by the sharp increase in the minimum number of contributory years required before they can get an early retirement pension. Beltrametti (1997) shows that the social security wealth of the self-employed fell on average, after the reforms, by 18 per cent, while for public sector workers it fell by 36 per cent; for pensioners, however, the fall is limited only to 5 per cent, entirely attributable to the 1992 reform.

Attanasio and Brugiavini (1997), using micro data for the 1991–93 period, conclude that households of public sector employees actually increased their saving rate more than the other groups, and interpret this as evidence of a reaction to the 1992 pension reform. Here we extend their analysis to other demographic groups and to more recent sample surveys, which may also take account of possible effects of the more recent 1995–97 pension reform.

The data To test the effects of pension reforms on private saving, we use the same six Bank of Italy Surveys of Household Income and Wealth that we used previously. We assume that the 1989 and 1991 surveys provide information about household saving behaviour before the reforms, and that the successive surveys should incorporate the possible reactions by the households to the reforms themselves. In the course of the analysis, we also consider the intermediate passages from some pairs of years, to check the sensitivity (and variability) of the results to the particular survey used. In each survey, households have been classified into different groups, according to the seniority (contribution years) of the head and his/her occupation. Then, the analysis proceeds to check whether these various groups changed their saving behaviour in a significantly different way as a consequence of the reforms. We have classified households in four groups, on the basis of the occupational position of the head: (i) private sector employee; (ii) public sector employee; (iii) self-employed; and (iv) other (that is, pensioner aged at least 60 years in 1989, 62 in 1991 and so on). Then, each of the first three groups has been divided into two subgroups ('young' and 'old'), according to the seniority of the head (less or more than 14 years of work in 1991, 18 in 1995 and so on). Among these categories, the second and third groups (and in particular the young) should have been more severely hit by the reforms, which reduced only marginally the social security wealth of those belonging to the fourth group. Table 7.8 reports some characteristics of the households in the various groups, in the first and last years of the sample period.

The size of the groups of 'young' workers tends to increase over the period, because they include all those workers with less than 15 years of seniority in 1992; in other words, this group is fed by the new entrants to the labour market in the years following 1992, independently of their age. The last group, 'others', decreases since it does not include those 'old' workers who reached retirement age between 1991 and 1998, in order to avoid the saving behaviour of this control group being an average of persons affected differently by the reforms. As expected, young households tend to be more educated and to have fewer members and income recipients. The average age of the 'young' groups increases more than that of the other groups, as a result of the increase in their size. The saving rate shows no change for young private employees and an increase for other employees, both private and public. The table does not enable us to detect whether the increase in saving of the 'young' group is the result either of the reforms or of the different changing features of the group.

After the exclusion of some households with implausible values of income or consumption (less than €2000 per year), the saving rate has been defined for each household as the difference between the natural logarithms of income and consumption. The use of logarithms has the effect of reducing the impact of outliers on the averages and on regressions results.

Table 7.8 Sample means for selected variables, SHIW 1989 and 2000

	Private sector employees				Public sector employees				Self-employed				Other	
	Young		Old		Young		Old		Young		Old			
	1989	2000	1989	2000	1989	2000	1989	2000	1989	2000	1989	2000	1989	2000
% Freq.	6.30	11.50	20.10	7.70	4.30	7.80	10.20	4.90	4.30	8.50	15.50	6.50	32.50	18.50
Age	30.10	37.10	46.40	49.20	31.60	40.50	47.30	51.00	31.30	39.90	50.90	52.30	70.30	78.70
Education (years)	10.70	10.60	8.40	9.50	12.80	12.90	10.70	11.40	12.00	12.00	8.10	9.50	5.90	5.90
N members	2.90	3.10	3.70	3.30	2.90	3.00	3.50	3.20	2.70	3.00	3.40	3.30	2.00	1.70
N earners	1.60	1.70	1.80	1.90	1.50	1.60	1.70	1.90	1.50	1.60	1.80	2.00	1.60	1.40
Income	24.19	25.91	27.94	31.64	25.76	26.54	29.51	31.67	28.02	31.67	33.87	36.68	17.26	17.23
Consumpt.	18.52	19.38	19.78	22.67	19.47	19.89	21.11	21.64	21.65	23.17	23.63	25.00	12.29	12.34
Saving rate	0.23	0.25	0.29	0.28	0.24	0.25	0.28	0.32	0.23	0.27	0.30	0.32	0.29	0.28

Source: Our computations from SHIW data; monetary values refer to €000 at 1998 prices. Individual characteristics refer to the head of the household.

The difference-in-differences estimation The first kind of evidence that we provide is based on a simple difference-in-differences approach, which requires the subdivision of the sample into a 'treatment' group and a 'control' group. The treatment group is composed of those units who are subject to an exogenous policy change. The control group includes those households not interested in the policy.

Let us consider observations relative to the periods before and after the policy change, and compute the differences in the values of the variable of interest for each group. In so doing, we are testing the effect of the policy by observing whether the change relating to the treatment group is significantly different from that observed for the control households. For the reasons already indicated, the control group of our analysis is the category 'other', essentially formed by those households whose head was already retired before the reforms took place. We have specified several treatment groups, described in the previous section, to observe any possible change in behaviour by households affected differently by the reforms. This method does not need the computation of social security wealth and its changes, but relies crucially on an appropriate subdivision of the observations between treatment and control groups.

If S_{it} is the variable of interest observed at time t for group i, it is possible to estimate this regression:

$$S_{it} = \alpha + \beta\, D_T + \gamma\, D_{POST} + \delta\, D_T D_{POST} + u_{it},$$

where the index i refers to the treatment or control group, the index t to the initial or final period, D_T is a dummy equal to 1 if the household belongs to the treatment group, and D_{POST} a dummy for the post-policy (final) period, reflecting aggregate shocks. The coefficient δ provides the estimate of the differential impact of the policy on the saving rate of the treatment group, separating the effect of the policy change from other unobserved determinants that were possibly changing over the same period.

The correct application of the difference-in-differences procedure requires some identifying assumptions. Among them, the most important is that the only significant difference between the treatment and the control groups consists in being or not being subject to the policy measure. Otherwise, the change in the saving behaviour of the different groups between the initial and final period might have been driven by other unobserved factors. This is clearly not our case, since the control group is significantly different from the others, so a simple application of the approach through the computation of means of the saving rates and their differences would be meaningless. We turn therefore to a regression-based approach, where the explicit introduction of demographic attributes among the regressors allows us to isolate the net effect of the change in the policy regime. Since the saving rate typically follows a hump-shaped profile over the life cycle, the systematic difference between the stages in the life cycle of the treatment and control groups can be neutralised with a polynomial in age, which should also take into account the change in the average age within each group between the beginning and the end of the sample period. The results presented on the effect of pension reforms on saving rates are not affected by the life-cycle profile of the saving rate.

Other systematic differences are apparent for average education and for the average size of the household. Our approach consists in a regression similar to the formula shown above, enriched with a vector of demographic variables, to isolate the net effect of the pension reform on the saving rate. The set of demographic factors that could influence the saving rate contains a third-order polynomial in age, the number of household members and of earners, and dummy variables for the sex of the head, his/her education level and the geographic area of residence.

Table 7.9 contains the basic results of this section; it reports the estimates for various subperiods of the coefficient δ, which should represent the net effect of the pension reforms on the saving rate of the various groups. For example, the coefficient 0.036 in the top left part of the table means that the difference between the changes in the saving rate for the group formed by

Table 7.9 Estimated coefficients of the interaction term 'post & treatment'

Treatment group	1989–93	1989+ 91–93	1989–2000	1989+ 91–2000	1989+91– 1998+2000
Private sector employees					
All	0.036*	0.059**	−0.051**	−0.028	−0.021
Young	0.043	0.063**	−0.055*	−0.043	−0.034
Old	0.031	0.053**	−0.064**	−0.038	−0.036*
Public sector employees					
All	0.068**	0.083**	−0.023	−0.008	0.008
Young	0.049	0.058*	−0.054	−0.055	−0.028
Old	0.076**	0.091**	0.021	0.031	0.039*
Self-employed					
All	0.003	−0.006	−0.026	−0.031	−0.026
Young	0.006	−0.028	−0.075*	−0.093**	−0.075**
Old	−0.001	−0.002	−0.031	−0.041	−0.020

Notes: OLS with robust standard errors; ** significant at 5%; * significant at 10%. Each regression includes a polynomial in age, number of components and earners, dummies for sex of head, his/her education level, and the area of residence.

Source: Our computations on SHIW data.

all private sector employees and for the control group is equal to 0.036 points of the saving rate. The estimates have been obtained through a set of regressions on the same control group and the various treatment groups, pooling the data for the years before and after the reforms. The estimation method is ordinary least squares, with White robust standard errors.[12]

As discussed above, a priori we should expect an increase in the saving rate by those groups which have been more severely affected by the reforms, in particular the 'young', those employed in the public sector, and the self-employed. We considered various pairs of years to test the sensitivity of the results to the choice of a particular final period.

The regressions over the period running from 1989 to 1993 should provide information about the effects of the first pension reform, implemented at the end of 1992. To select a more stable initial period, we have also pooled together the first two available surveys, 1989 and 1991, and compared them with the 1993 survey. Both private and public sector employees increased their saving rates in this period more significantly than the self-employed, with respect to the treatment group:[13] the change is substantially similar for the young and the old, and greater for public sector employees. These results could actually be interpreted as evidence of a sharp reaction to the reform (see also Attanasio and Brugiavini 1997). However, during the same period, after the currency crisis, a much tighter budget policy was enacted, which might have signalled a general reduction in the current and future private income net of taxes and transfers, which, in its turn, might explain the rapid diffusion of the signal.

Considering more recent surveys, the evidence of an increase in the saving rate by most demographic groups becomes definitely weaker. Taking into consideration the whole available period, from 1989 to 2000 (or from 1989 *and* 1991 to 2000), no treatment group shows a relative increase in its saving rate, on the contrary in general the decreases are more frequent, in particular for the private sector employees. The results do not change substantially if we pool together the first two surveys to form the initial period and the last two surveys to represent the final period (last column of the table). The only significant increase that remains in this case is that shown by the older civil servants, while, at first sight, we should expect a stronger reaction from the younger. Actually, the real change in the civil servant pension system mainly consists of the abolition of the very generous possibility of early retirement benefit, usually combined with new working activity in the hidden economy. While younger civil servants had already excluded this possibility from their future, for the older ones it really was an unexpected change in their life plan. The young independent workers, those on the whole more severely affected by the reforms, actually show a significant relative decline in their saving rate.

These results could be influenced by the different levels of permanent income and living standards of households belonging to the various treatment groups. In the same demographic group, more wealthy households, due to a higher level of sophistication in their financial decisions and a better knowledge of the implications of the policy changes, could offset the shock to a larger extent than other households could. Current disposable income cannot be used to classify households into different levels of wellbeing, since income is correlated with the saving rate; the education level of the household head is thus used here to approximate the distribution of permanent income across households; we distinguish three levels of education, primary (from 5 to 8 years of schooling), secondary (at least 13 years) and tertiary (at least 17 years).

Table 7.10 shows the estimated coefficients of the interaction term between the treatment group dummy and the final-period dummy for smaller treatment groups, one for each educational category. In general, the relative rise in the saving rate between 1989 and 1993 appears to be due mainly to the behaviour of the more-educated households. This is particularly apparent for private sector workers, both employees and self-

Table 7.10 Estimated coefficients of the interaction term 'post & treatment' by education group, 1989+91–93 versus 1989+91–2000

	1989+91–93			1989+91–2000		
	Primary	Secondary	Tertiary	Primary	Secondary	Tertiary
Private sector employees						
All	0.040*	0.099**	0.177**	−0.054**	−0.024	0.052
Young	0.029	0.117**	0.120	−0.026	−0.051	−0.064
Old	0.042*	0.078**	0.220**	−0.071**	−0.0113	0.208**
Public sector employees						
All	0.055*	0.123**	0.097**	−0.021	0.002	−0.018
Young	−0.058	0.109**	0.116**	−0.128**	−0.060	−0.006
Old	0.091**	0.114**	0.076	0.036	0.050	−0.021
Self-employed						
All	−0.029	0.113	0.122**	0.036	0.050	−0.021
Young	−0.155*	0.078	0.076	−0.034	−0.022	−0.078
Old	0.0006	−0.068	0.215**	−0.185**	−0.028	−0.109

Notes: OLS with robust standard errors; ** significant at 5%; * significant at 10%. Each regression includes a polynomial in age, number of components and earners, dummies for sex of head, his/her education level, and the area of residence.

Source: Our computations on SHIW data.

employed. Also in the passage from 1989 + 1991 to 2000, the relative change in the saving rate is generally greater for the more- than for the less-educated households, who in many cases experienced a relative decline in the saving rate.

Aggregate numbers therefore conceal some reactions by specific demographic groups. The greater reaction by the more educated could be due not only to a greater ability to foresee the future consequences of current policy changes, but also to a change in the distribution of income favouring more-skilled workers; there are some indications that in the 1990s the distribution of income has actually become more polarised in favour of the more educated, who therefore were able to increase their relative saving rate. The introduction of a contribution-related pension scheme, which reduces the advantage of dynamic careers, may also have played a role in the saving behaviour of the more educated.

The results suggest therefore that a significant differential change in the propensity to save can be observed for most groups in the first part of the period, and this could be interpreted as an immediate reaction to the first pension reform.[14] In the following period, always keeping fixed the initial period as reference for comparisons, the only significant change that persists concerns the older public sector employees. These estimates could be interpreted as showing that the possible reduction in pension wealth produced by the reforms had only a limited effect on private wealth accumulation (see also Jappelli 1995), and that the consistent reaction of saving is concentrated in particular among the more-educated and richer households. Over the whole 1989–2000 period, the relative increase in the saving rate within each treatment group is generally larger for the more-educated households. The results obtained so far seem also to confirm the presence of negative cohort effects on saving for the younger generations (relatively lower for the more educated), already pointed out in previous sections of the chapter.

Summary

We can draw the following preliminary conclusions:

- the 1991–93 reactions of the treatment groups show that the dramatic message of the financial crisis produces a relative increase in the saving rate (or a lower decline) of the groups affected by the reform;
- in the medium run, when general financial and currency crisis have been reabsorbed, the impact on saving rates of the groups affected by the reforms becomes weaker;
- within the treatment groups, the analysis shows that in the medium run the relative increase in the saving rate remains strong either if the level of education is higher or if income is higher (the two are not independently distributed);

- in the case of lower-income households, they seem to behave as if the dramatic message of the financial crisis produced a temporary reduction of their impatience *vis-à-vis* the future;
- although we have not yet computed the substitution rates of private to social security wealth, the results we have obtained at this stage suggest a trivial statement: they may depend on the amount of information held by households, which in its turn depends on their capability of being informed.

Microeconomic Effects of the Reforms: Individual Social Security Wealth

The analysis of the previous section was based on qualitative evaluations of the change in social security wealth for different household groups. In order to evaluate the quantitative importance of the difference-in-differences estimates we should compare the realised change in private wealth accumulation with the actual reduction in social security wealth. In this section, the analysis of the effects of pension reforms on saving will be extended to take into account the quantitative evaluation of social security wealth for each household of the sample before and after the reforms. The quantitative changes of the social security wealth will be compared with changes in saving and private wealth holding across time.

Computation of social security wealth

The survey (SHIW) data do not contain information on social security wealth. In this subsection we describe how we have estimated this variable. Social security wealth is defined as the sum of future expected benefits net of the sum of future expected contributions. According to this definition, at each point of time and for each individual, social security wealth expresses the accrued entitlements to wealth which the individual has realised through his/her participation in the public pension scheme.

For each individual observed in the six surveys (1989, 1991, 1993, 1995, 1998 and 2000), first, we compute the present value of future pension benefits. In so doing, we use information on age, sex, occupation, seniority, expected retirement age, lifetime earnings profile, life expectancy and the relevant social security legislation in the year of observation. Next, we estimate the present value of future contributions to the pension scheme that an individual belongs to. Finally, we net out the present value of contributions from the present value of benefits to obtain the expected net social security wealth for each individual at the time of the survey. In other words, the net social security wealth ($SSWN$) is the difference between the present value of future benefits (PVB) and the present value of future contributions (PVC) evaluated at the time of observation for each individual in the sample.

The formula of the social security wealth for an individual i, who belongs to a j pension scheme, at time t is defined as:

$$SSWN_{i,j,} = \sum_{k=L+1}^{E(T)} \frac{P_{i,j,k}}{(1+r)^{k-t}} - \sum_{k=t}^{L} \frac{a_{i,j,k} W_{i,j,k}}{(1+r)^{k}}, \quad (7.1)$$

where:

$P_{i,j,k}$ = pension benefit of individual i who belongs to the j pension scheme, measured at time k;
$a_{i,j,k}$ = contribution rate for individual i who belongs to the j pension scheme, measured at time k;
$W_{i,j,k}$ = gross wage of individual i who belongs to the j pension scheme, measured at time k;
r = discount rate;
L = expected age of retirement;
$E_L(T)$ = life expectancy at age L;

The computation of equation (7.1) produces a value of the net social security wealth for each worker in the survey. As a whole, after considering individuals with a full career (employees and self-employed) with gross annual earnings higher than €4000 and lower that €100000, in order to exclude outliers from the computation, we have 30078 observations for employees and 6010 observations for the self-employed. The following hypotheses are used in the computation of social security wealth:

1. we express all values in 1998 constant prices and we assume perfect foresight about future inflation and a complete price indexation mechanism;
2. we assume that when an individual retires, he/she has a good knowledge of pension rules and he/she will compute future taxes and benefits as if the pension legislation will persist in the future, unless new information on pension legislation becomes available to individuals;
3. the age of retirement is taken for each individual from subjective expectations expressed by respondents in the surveys;
4. life expectancies are taken from official statistics of the Italian Statistical Bureau in 1996;
5. lifetime earnings profiles that are used to compute future pension benefits and contributions are estimated for three different level of education, ten cohorts, and for employees and self-employed separately (a complete description of the estimation procedure of lifetime earnings profiles is given in the next subsection);

6. the discount rate used to compute the present value of pension benefits and contributions is assumed to be constant and equal to 3 per cent.

Estimation of lifetime earnings profiles We estimate lifetime earnings profiles for sex, occupation and educational level. We consider dependent and self-employed workers; and three levels of education: primary, secondary and tertiary. The combination of these economic and demographic characteristics gives us 12 stylised lifetime profiles. Using information on birth date and on estimated gross earnings of each individual we are able to generate 'individualised' earnings profiles, where the shape corresponds to one of the 12 stylised profiles, and where the level of gross earnings depends on the birth date of the individual to whom the profile refers.

Income figures in the surveys (SHIW) are net of personal income taxes and social security contributions. The first step in the estimation was the construction of gross earnings. We used the following procedure: define YN_i as the net income of an individual i; then his/her gross income is calculated by solving the following relation:

$$YL_i = \frac{[YN_i - (t_{j+1} - t_j)\ Y_{j+1} - (t_{j+2} - t_{j+1}) Y_{j+2} - \ldots - (t_{j+n} - t_{j+n-1}) Y_{j+n} - D_i]}{(1 - t_{j+n-1})} \tag{7.2}$$

where:

Y_j = income bracket of the personal income tax;
t_j = marginal tax rate on income bracket j;
D_j = tax credit;
$YL_{j+n} < YN_i < YL_{j+n-1}$.

Next, compute gross earnings for the 1989, 1991, 1993, 1995, 1998 and 2000 surveys, by taking account of the changes in the personal income tax law occurred through the period. In the computation of gross income only full-time workers were considered.

Earnings profiles by sex, occupation and education level are obtained by regressing the logarithm of gross earnings on a polynomial on age and a set of cohort dummies. Separate profiles for employees and self-employed, for men and women and for three levels of education were obtained. Two stylised profiles are described in the Figure 7.11 for male and female employees. For each observed individual in the surveys, the first gross earnings is computed by re-scaling the estimated earnings by both an age and a cohort parameter. In the same way, the earnings in the year before retirement is estimated by using information on expected retirement age. The individualised earnings profile was used to estimate the average growth of gross salary and the seniority of each individual at the age of retirement. This information is then used in the computation of the social security wealth for each individual.

The different shape of the earnings profiles can be observed in Figure 7.11. They reach a maximum before the age of retirement for low-skilled workers, whereas the slope of the profile is positive throughout the working life period both for the second and third education levels (secondary and tertiary).

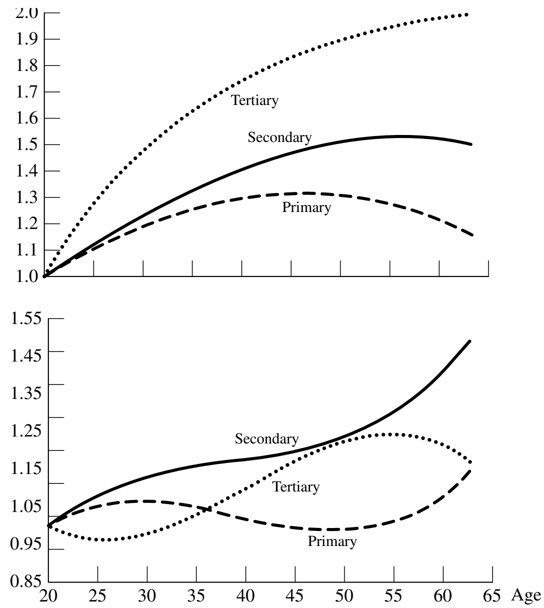


Figure 7.11 Lifetime earnings profiles for male (top) and female (bottom) employees and three different levels of education

Effects of the 1992 and 1995 pension reforms on individual social security wealth in the surveys (SHIW) During the period under observation, three reforms of the pension system occurred. The main legislative changes of these are reported in detail in Section 2 (see also Franco 2001). This subsection deals with the details of the computation of social security wealth in 1989 and 1991, the years before the first reform, and in 1993, 1995, 1998 and 2000, when the reforms had already been undertaken.

SOCIAL SECURITY WEALTH IN THE 1989 AND 1991 SURVEYS

1. Employees in the survey compute their pension benefits according to the pre-reform earnings-related formula:

$$P_i = c^* N_i^* 0.2^* W_{i,1}^* (1 + m_{i,j})^{Ni-5*} \left[\frac{(1 + m_{i,j})5 - 1}{m_{i,j}} \right]; \qquad (7.3)$$

the self-employed compute their pension benefit according to:

$$P_i = c^* N_i^* 0.1^* W_{i,1}^* (1 + m_{i,j})^{Ni-10*} \left[\frac{(1 + m_{i,j})10 - 1}{m_{i,j}} \right]; \qquad (7.3')$$

where:

c = 0.02 is the internal rate of return of the earnings-related formula;
N_i = seniority at retirement for individual *i*;
m_{ij} = average growth rate of real gross wage for individual *i* and pension scheme *j*;
W_{i1} = gross wage in the first year of work of individual *i*.

Values of W_{it}, N_i and $m_{i,j}$ are estimated following the procedure described above. They reflect the combination of personal characteristics such as the individualised average growth of gross earnings, the seniority at retirement, the level of earnings and the shape of one of the stylised gross earnings profiles.

2. All pension benefits are indexed to real wages growth, which we fixed at 1.5 per cent per year.
3. Contributions rate is equal to 27.4 per cent and 12 per cent for employees and the self-employed, respectively.

SOCIAL SECURITY WEALTH IN THE 1993 SURVEY

1. Individuals were divided according to their seniority. Define seniority in 1993 of individual *i* with $S_{i,1993}$, then three groups of individuals are identified: (i) $S_{i,1993} \geq 16$ years; (ii) $S_{i,1993} < 1$, and (iii) $1 < S_{i,1993} < 16$. Individuals who belong to group (i) do not change the formula of computation of the pension benefits with respect to the formula used to compute the pension benefit in 1991. Individuals who belong to group (ii) compute their pension benefits according to the new pension formula introduced in 1993 by the Amato government:

$$P_i = a\left(\frac{1}{N_i}\right)\left\{ W_{i,1}\left[\frac{(1 + m_{i,j})^N - 1}{m_{i,j}}\right] + 0.01\sum_{k=1}^{l-1} W_k (1 + m)^{1-k} (L - k). \right. \qquad (7.4)$$

Individuals who belong to group (iii) compute their pension benefits according to the so-called pro rata rule, that is, their pension is calculated according to the 1991 rule for years before 1992 and according to the new rule introduced afterwards.

2. The lifetime profile of pension benefits is constant in real terms for each individual, that is, pensions are no longer indexed to real wage growth.
3. Retirement age is increased according to subjective expectations expressed by individuals in the survey.

SOCIAL SECURITY WEALTH IN THE 1995, 1998 AND 2000 SURVEYS

1. Again, individuals observed in the two surveys are divided according to their seniority. In 1995 the three groups are: (i) $S_{i,1995} \geq 18$ years; (ii) $S_{i,1995} < 3$ and (iii) $3 < S_{i,1995} < 18$. In 1998, individuals are divided according to the following rule: (i) $S_{i,1998} \geq 21$ years; (ii) $S_{i,1998} < 5$ and (iii) $5 < S_{i,1998} < 21$ years. In 2000, individuals are divided according to the following rule: (i) $S_{i,2000} > 23$ years; (ii) $S_{i,2000} < 7$ years; and (iii) $7 < S_{i,2000} < 23$ years.

   Pension benefits of individuals who belong to group (i) in 1995, 1998 and 2000 are computed according to the 1991 pension rule. Pension benefits of individuals who belong to group (ii) are calculated using the notional DC rule introduced by the Dini government in 1995:

$$P_i = k^* a_{i,j}{}^* W_{i,1} \left[\frac{(1+g)^{Ni} - (1+m_{i,j})^{Ni}}{g - m_{i,j}} \right], \tag{7.5}$$

   where:
   g = 1.5% is the real growth rate of GDP;
   m_{ij} = real growth of individual gross wage;
   a_{ij} = 33% for employees and 20% for the self-employed;
   k = coefficient of conversion provided by the pension law which depends on the average expected life at retirement and on a 1.5% discount rate.

   Pension benefits for individuals who belong to group (iii) are calculated according to the pro rata rule, that is, their pension is computed according to the three rules described above for years before 1992, from 1992 to 1995 and after 1995, respectively.
2. Contributions rate for employees is raised from 27.4 per cent to 32.7 per cent since 1995 onwards.
3. Contributions rate for the self-employed is raised from 12 per cent to 15 per cent in 1995; from 15 per cent to 19 per cent since 1998 onwards.

A crucial variable in the computation of social security wealth is the retirement age. In particular, when the pension formula is earnings related there is an incentive to retire earlier because the amount of the benefit is not directly related to life expectancy at retirement. In the computation of net social security wealth we decided to consider retirement age as it is reported by the subjective expectations expressed by each individual in the survey. A feasible alternative would have been the imputation of the 'legal' retirement age in each year of the survey to each individual. Both the first and the second alternatives have advantages and disadvantages. We preferred the first, because it appears more useful for one of the main aims of the chapter, that is, the estimation of the degree of substitutability between public pension and private saving. By considering the subjective expectations of the retirement age we are able to capture the individual expected working time from the year of observation onwards. This variable gives us a more precise picture of the planned future earnings for each individual, a crucial variable for the determination of the optimal level of consumption and saving in an intertemporal context.[15]

Analysis of net social security wealth and its relation with total private wealth

In this subsection we examine the path of social security wealth and some other indicators of the expected values of variables in the social security system, as they result from the micro data of the six surveys used in this chapter, before and after the structural reforms of the last decade.

Both stock and flow variables are considered. The stock variables are: the present value of pension contributions (*PVC*), the present value of pension benefits (*PVB*), and the resulting net social security wealth (*SSWN*), as defined in equation (7.1). Pension benefits are calculated in each year of observation by applying equations (7.3), (7.3′), (7.4) and (7.5) or the appropriate combinations of them according to the seniority of the individual at the time of observation. Gross earnings are estimated following the procedure described above (pp. 291–2), and are then used for the computation of pension benefits and contributions. Each stock variable is measured at time *t*, therefore the value of *PVB*, *PVC* and *SSWN* depends on the expected level of contributions, benefit and discount rate; moreover, it depends on the age of the individual at the time the variable is computed. In order to have an index which measures the pension benefit in terms of life-cycle resources, hence independent of the age of the individual when he/she is observed, for each individual belonging to the sample, we divide the annuity value of the pension benefit by the annuity value of earnings, both evaluated at the year of retirement. This ratio supplies a *time-independent* measure of the replacement rate of pension resources to lifetime earnings.

Other indicators considered are: the expected age of retirement; the expected replacement rate defined as the ratio of the first gross pension benefit to the last gross earnings for each individual in the survey *RR*; and the expected seniority at retirement *SR* defined as the estimated expected total number of years of contributions to a pension scheme by each individual in the survey. These may be useful for analysing some microeconomic effects of the reforms, for example, on labour supply.

In Table 7.11, we report some information about the expected effects of the reforms of the 1990s in terms of these indicators. We consider in particular the expected value of the first-year pension benefit, the average replacement rate and the average replacement rate for the young (<15 years of seniority in 1992) and adult (>15 years of seniority in 1992) workers, and the expected seniority at retirement. Values of the pension benefits seem to signal a general tendency to decrease. However, they can be misleading because they are computed on different samples. Other variables give more interesting indications. The average expected replacement rate declines slightly for employees. The reduction is stronger for those self-employed whose benefit was reduced significantly by the introduction of the notional DC scheme in 1995 and afterwards.[16] It is interesting to split the population into young and adult workers. If we look at the data reported in the third and fourth rows, we see that the reduction in the replacement rate is stronger for young workers, in particular if they are self-employed. This difference can be explained by the slow speed with which the pension formulae have been introduced for adult workers both in the 1993 and 1995 reforms. Expected seniority at retirement increases on average for employees, whereas it remains constant for the self-employed.

A similar effect is noticeable in Table 7.12, where we report the expected retirement age for different groups of the population. The increase is on average equal to 1.6 years for employees and 1 year for the self-employed. Young workers seem to expect a longer stay in the labour market compared with adult workers, perhaps as a reaction to the more severe computation rules introduced in the pension system. There is also a common tendency to raise the expected retirement age for different levels of education.

Expectations about retirement age do not appear to be very different if we split the data by cohort. Figure 7.12 displays the average value of expected retirement age from 1991 to 2000 for different cohorts. There is a common tendency, after 1993, to increase the expected number of working years for young and adult generations. At the end of the observation period expected retirement age is higher for young workers among employees and it is higher for adult workers among the self-employed.

Other important effects of the reforms are outlined by the total amount of resources which individuals in the sample were expecting in different years. Table 7.13 shows the average values of *PVC*, *PVB* and *SSWN* for employees and the self-employed in selected years. The last row of the table, where the average net social security wealth is reported, shows the general strong reduction in the net liabilities of the social security system, after the

Table 7.11 Expected value of pension benefits, replacement rate and seniority at retirement for employees and the self-employed

	1991		1993		1995		2000	
	E	SE	E	SE	E	SE	E	SE
Pension benefits (€)	16167	18864	16211	16590	15547	13911	16345	14395
Replacement Rate (%)	71.4	71.9	67.6	68.5	68.2	58.6	69.8	56.2
RR for adult worker (%)	–	–	69.2	69.0	74.4	75.9	75.8	75.7
RR for young worker (%)	–	–	65.9	68.0	63.1	46.8	67.4	46.5
Seniority at retirement (yrs)	36.6	38.8	36.7	37.8	37.6	38.7	38.7	38.4
Income (€)	16815	18856	18216	16580	17065	17303	17866	21745

Notes: Each number represents the value expected by the average employee (E) and/or self-employed person (SE) at work in year t, $t = 1991$, 1993, 1995 and 2000. Adult/young worker is defined as an individual with seniority allowing him/her to belong to the first/third group described above (pp. 292–5).

Table 7.12 Subjective expectations of retirement age, 1991 and 2000

	1991	2000
Employees	59.3	61.0
Young workers	59.2	61.7
Adult workers	59.9	60.1
Self-employed	61.4	62.4
Primary	59.2	60.7
Secondary	58.9	61.5
Tertiary	60.2	62.2

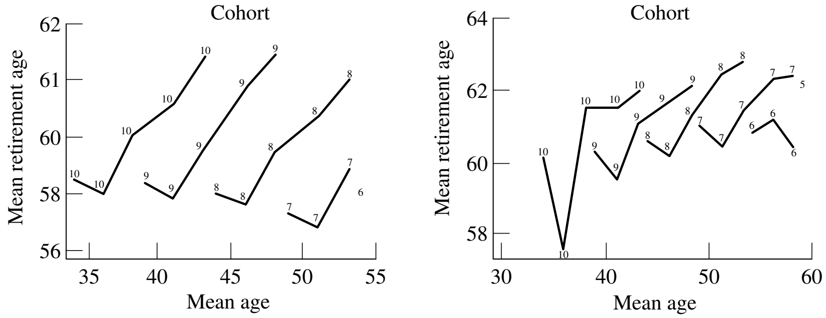


Figure 7.12 Expected retirement age by cohort of employees (left) and self-employed (right)

Table 7.13 Expected present value of contributions, pension benefits and net social security wealth for employees (E) and self-employed (SE), 1991–2000 (€ constant 1998)

	1991		1993		1995		2000	
	E	SE	E	SE	E	SE	E	SE
PVC	75101	39289	93059	57010	82958	53651	95305	61.604
PVB	157545	67335	135558	31392	129269	101188	128237	17663
SSWN	82443	128045	42498	74381	40001	47536	32931	56.058

reforms of the 1990s. In absolute terms, the 1992 reform appears the most effective, whereas the 1995 reform did not seem to reduce *SSWN* significantly, at least on average. The two components of *SSWN* both contribute to its reduction.

The *PVC* increases and the *PVB* decreases; this is true to a larger extent for the self-employed. The average value of the net pension multiplied by the number of employees and self-employed wealth can be used to calculate the total amount of pension liabilities of the social security system. According to this very approximate computation, total net pension liabilities from 1991 to 2000 were reduced by 52 per cent. The ratio of net social security wealth of active cohorts to GDP at 1998 prices decreased from 201 per cent in 1991 to 96 per cent in 2000. It is not easy to compare these calculations with other more sophisticated estimations of pension liabilities. First, we do not consider the pension entitlements of retirees; second, we do not take explicitly into account either survivor or invalidity pensions; finally, the discount rate used in the computation of the net social security wealth (3 per cent) may not be the same as is used in other estimations. Given all these differences, however, the magnitude of the effects appears to be comparable with those of other studies (Van Der Nord and Herd 1993; Beltrametti 1995; Rostagno 1996; Zollino 2001).

The reduction of *SSWN* is unevenly distributed among individuals in the surveys. Moreover it depends on more measures. The most effective one in the reduction of *PVB* is the general abolition of the indexation mechanism introduced in 1993. Changes in the eligibility and computation rules have different effects on individuals in the sample. As we stressed above, the transition to the contribution-related pension scheme is very slow, especially for adult workers. Table 7.11 showed that the expected replacement rate will be almost unchanged for workers with more than 15 years of seniority in 1992. The reforms are expected to have stronger effects on the self-employed than on employees, because of the introduction, in 1995, of a contribution-related pension scheme coupled with the lower contribution rate of this category (15–20 per cent instead of 32.7 per cent). An increase in the contribution rate also has an effect on *SSWN* through greater values of *PVC*. A more controversial effect on *SSWN* is the one exerted by the increase in the retirement age because it determines an increase in the value of future contributions and in the value of future benefits. The relative strength of the two effects determines the sign of the change in *SSWN*. Table 7.14 summarises other distributive features of the changes in the average values of *PVC*, *PVB* and *SSWN*.

The reduction in the net social security wealth is stronger for employees (60.1 per cent), workers with less education (50.9and 58.2 per cent) and women (54.7 per cent). *PVC* rises, especially for the self-employed (56.8 per

Table 7.14 Changes in the present value of contributions, benefits and net social security wealth, by occupation, education level and sex, 1991–2000 (%)

	Employees	Self-employed	Primary	Secondary	Tertiary	Men	Women
PVC	26.9	56.8	22.0	19.0	31.1	30.2	21.7
PVB	−18.6	−29.7	−23.0	−21.4	−9.6	−19.4	−23.3
SSWN	−60.1	−56.2	−50.9	−58.2	−45.0	−52.8	−54.7

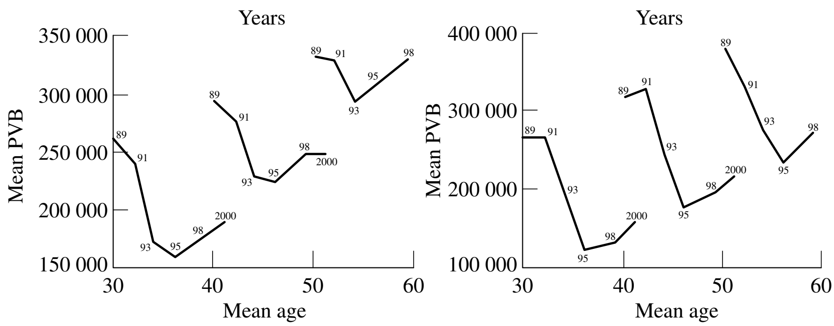


Figure 7.13 PVB *by cohorts and age: employees (left) and self-employed (right)*

cent) and for well-educated workers (31.1 per cent); whereas *PVB* decreases strongly for the self-employed (−29.7 per cent), women (−23.3 per cent) and less-educated workers (−23.0 per cent). Note that a reduction in *SSWN* can have different effects on saving if it derives from a reduction in the present value of future benefits or from a rise in the present value of future pension contributions. In the second case, account should be taken of the presence of liquidity constraints for individuas affected by the reform.

Further evidence of the different effects of the pension reforms by age is given in Figure 7.13 where the value of *PVB* is shown by ten-year cohorts and occupation. The reduction in the future expected value of *PVB* is particularly strong for the younger generations and for the self-employed. In particular, the absolute value of PVB decreases for all the average cohort values in 1993. The reduction in the present value of future benefits continues in 1995 for the self-employed. From 1995, the value of *PVB* increases for each cohort. However, we must take into account the fact that the paths of *SSWN* and its components (*PVC* and *PVB*) are driven both by changes in the pension rules, and by time.

Table 7.15 Percentage change in the ratio between the annuity value of SSWN *and the annuity value of all lifetime earnings, both evaluated at the age of retirement*

	1993–91		1995–93		1998–95	
Age in 1991	E	SE	E	SE	E	SE
46	−16.44	−15.52	2.30	−10.64	−7.44	12.76
41	−17.91	−18.41	9.75	−12.18	−2.03	11.23
36	−18.20	−13.92	7.13	−13.35	−0.49	1.91
31	−21.39	−18.32	7.01	−13.78	−0.84	−1.83
26	−23.47	−19.89	3.85	−29.90	−2.81	1.41
21	−27.68	−23.48	−0.09	−38.57	−2.18	1.07
16	−31.08	−32.31	0.10	−33.60	−4.93	−4.06

Note: E = employees; SE = self-employed.

As time passes, the date of retirement gets closer for each cohort and this fact alone makes the value of *PVB* larger (at least, until the age of retirement) and the value of *PVC* smaller (fewer contributions to pay). Therefore it might be misleading to compare the value of *PVB* or *SSWN* for different cohorts in a certain year. In order to disentangle the first effect (changes in *SSWN* caused by the reform) from the second one (time effect) we use the ratio between the annuity value of *SSWN* and the annuity value of the discounted value of the whole working life earnings, both evaluated at the age of retirement. This ratio gives us a 'time free' measure of the effects of the reforms on the net social security wealth of individuals of a different age. Table 7.15 reports the percentage change of this ratio between the two years indicated. If we look at the table by column, we have a measure of the effect of pension reforms in a specific year for different cohorts. There seems to be a common tendency: in almost all cases, the change in the ratio is more pronounced for younger generations, even though the reduction appears to be relevant, in particular from 1991 to 1993, for all generations. If we look at the table by row we have a measure of the different effects of the reforms in the decade. In this case, the stronger effect of the 1992 reform becomes quite evident. Later, the effects show a greater contrast: from 1993 to 1995 the reduction in the ratio is more pronounced for the self-employed; from 1995 to 1998 it is more pronounced for employees.

Pension and private wealth: household behaviour

The degree of substitutability between social security wealth and total private net wealth is the objective of the next part of the project. First, we show how these two variables are measured and computed in the six surveys. We aggregate both social security wealth and total net private wealth by household. In particular, we consider only households where at least one member is working. Social security wealth is now defined as the sum of the husband and wife's social security wealth. Total private wealth (*PW*) is the sum of financial and real assets less the sum of financial liabilities for each household.

Table 7.16 reports some statistics for the two variables in the six observed years. The mean value of net social security wealth by family decreases from an average value of €150 thousand in 1989 to a value of €70 thousand in 2000, that is, a reduction of 53 per cent. During the same period, total private wealth increases from €104 thousand in 1989 to €161 thousand in 2000, that is, an increase of 54 per cent. By comparing mean and median values for *SSWN* and *PW*, we note that net social security wealth is more equally distributed than private wealth, a result common to other developed countries. The ratio of the mean to the median value for private wealth decreases from 1.7 to 1.59 from 1989 to 2000, whereas the same ratio increases from 1.19 to 1.72, in the same period, for net social security wealth. From this point of view, the mitigating effect of the social security system on the uneven distribution of private wealth has been reduced by the reforms of the pension system.

Table 7.16 Mean and median values for household net social security wealth and private wealth in different years (constant 1998 €).

	SSW			PW		
Year	Mean (A)	Median (B)	A/B	Mean (A)	Median (B)	A/B
1989	150051	125770	1.19	104537	61084	1.71
1991	140449	122896	1.14	121532	80875	1.50
1993	84472	70070	1.21	154001	97354	1.58
1995	78327	58198	1.35	149820	99938	1.50
1998	74846	50309	1.49	156400	103808	1.51
2000	70255	40655	1.72	161808	101639	1.59

In Figure 7.14 we report the average household total wealth for active cohorts in the observed period. The effects of the reduction in social security wealth is more evident for the two younger cohorts. Older working cohorts seem to show a tendency to recover the values of total wealth in the years after the first reform. As the generation approaches retirement age, the value of total wealth increases because of the time effect (future pension benefits are less discounted).[17]

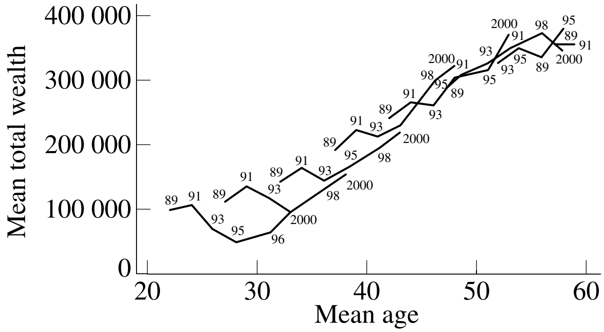


*Figure 7.14 Total household wealth (*SSWN + PW*), by cohort and year (average values, constant 1998 €)*

In the remaining tables we compare the situation just before the first reform with that in 2000. In Tables 7.17 and in 7.18 we can see that the reduction in the net social security wealth of families is more pronounced for those whose breadwinner has either a secondary or tertiary level of education and comes from the less affluent part of the country. In particular if we split the sample by geographical area, we have a confirmation of the effects of the reforms in terms of the distribution of total wealth. Families

Table 7.17 Mean values for household SSWN *and* PW *by different levels of education, 1991 and 2000 (constant 1998 €)*

	1991	2000	%	1991	2000	%
Primary	128 867	70 278	−45.5	97 053	114 129	17.6
Secondary	152 153	66 472	−56.3	137 533	176 388	28.3
Tertiary	168 974	77 510	−54.1	210 367	283 113	34.6

Table 7.18 Mean values for household SSWN *and* PW *by geographical area (constant 1998 €)*

	SSWN			*PW*		
	1991	2000	%	1991	2000	%
North	153 635	80 667	−47.5	119 850	189 477	58.1
Centre	142 913	63 994	−55.2	143 649	134 525	−6.4
South	119 885	54 533	−54.5	109 903	132 550	20.6

living in the South, the less affluent part of Italy, have 54.5 per cent less net social security wealth and only 20.6 per cent more total private wealth.

Finally, it is interesting to compare the distribution of net social security wealth and total private wealth by age and by occupation. In Table 7.19 we describe the evolution of the share of social security net wealth and private wealth. In 1991 the share of *SSWN* over total wealth (*TW*) was 59.7 per cent; in 2000 it decreases to a ratio of 38.6 per cent.

Table 7.20 reports the composition of the sum of total wealth in different years for employees and the self-employed. The composition of total wealth between the private and the public component seems balanced at the beginning of the period observed. With reforms in the pension system taking place during the decade, the ratio between private and total wealth increases, particularly for the self-employed.

Table 7.19 Composition of total household wealth in different years (%)

Year	*SSWN/TW*	*PW/TW*
1991	59.7	40.3
1993	24.7	75.2
1995	43.2	56.7
1998	42.0	58.0
2000	38.6	61.4

Table 7.20 Percentage of total household wealth in different years, by employees and self-employed

	1991		1993		1995		2000	
	SSWN	*PW*	*SSWN*	*PW*	*SSWN*	*PW*	*SSWN*	*PW*
Employees	55.0	45.0	40.1	59.9	41.2	58.8	41.5	58.5
Self-employed	44.8	55.2	33.2	66.8	24.5	75.5	25.4	74.6

The results presented in this subsection give a first indication of the problem of the substitutability between private and social security wealth. Retirement is not the only motivation for saving and wealth accumulation.

Precautionary saving, bequest motives and intergenerational transfer are some other important factors. We shall now present some econometric evidence for the existence of a causal relation between the path of social security net wealth and private wealth, which should give us a more precise measure of the degree of substitutability between these two variables.

The Attanasio–Gale correction

The analysis developed above has shown that the effects of the pension system reform have been different for different groups. In this subsection we use the computed social security wealth for families observed in the six surveys to analyse the degree of substitutability between private and social security wealth. A preliminary problem to solve is that the reaction of consumption and saving to a change in the future entitlement of social security wealth depends on the age of the individual at the time of the observation. This point was first raised in the pension-saving literature by Gale (1998) who suggests a method to correct the value of net social security wealth in the regression which estimates the substitutability between private and social security wealth. Consider, for example, two individuals who differ only with respect to their age. Within the life-cycle model, the same reduction in social security wealth occurring at time t will generate a different reaction in saving and wealth accumulation: the younger individual will have more time to absorb the reduction in social security wealth than the older. Therefore changes in consumption and saving at time t will be higher the older the individual. In order to take account of these problems we use a parameter which corrects social security wealth in the regressions. The parameter takes account both of the age effect and of the number of years that have elapsed from the current observation time and the time when the reform took place. In the construction of the correction parameter we follow the method suggested by Gale (1998) and which has recently been revisited by Attanasio and Rohwedder (2001).

The theoretical framework which supports the construction of such a parameter is the life-cycle model. Consider a situation without uncertainty where an individual works for L years and receives pension benefits for P years. The life span of the individual is certain and equal to $T = L + P$. In the first L years of his/her life, the individual works and accumulates savings, which are the sum of public pension contributions and private voluntary saving. The individual maximises a logarithmic utility function:

$$U = \sum_{t=1}^{T} \gamma^{t-1} \log C_t. \tag{7.6}$$

under the intertemporal budget constraint:

$$\sum_{t=1}^{T} C_t R^{t-1} \leq \sum_{t=1}^{L} (1-\alpha) W_t R^{t-1} + \sum_{t=L+1}^{T} B_t R^{t-1} \tag{7.7}$$

where:

γ is the factor which discounts future utility;
R is the factor of financial discount;

C_t is consumption at time (t), $t=1, 2, \ldots T$;
B_t is the pension benefit, $t=L+1, L+2, \ldots, T$;
W_t is gross earning at time t;
$WN_t = (1-\alpha)\ W_t$ is the net of contributions earning at time t;
$T=L+P$ is the lifetime horizon for consumption and saving planning;
α is the constant rate of contribution to the pay-go system.

In this very simple version of the model, the degree of substitutability between private and social security wealth should be equal to one, but the reaction of consumption and saving to the same change in future pension benefits is different when we consider different ages.[18]

From the solution of the optimisation problem we derive, for a given path of current disposable income WN_t, the optimal level of saving and private wealth accumulation PW_t at each time $t=1,\ldots, T$. In order to obtain a solution to the wealth accumulation equation that we shall estimate below, we first derive the value of PW_t in the absence of any reform in the pension system. In this case, we find a coefficient, which relates, at each age, $SSWN$ (net social security wealth) to PW (private wealth). Next, we consider a situation where at time k a reform of the public pension system takes place. In this case, the individual re-programmes his/her optimal path of saving and wealth accumulation from k onwards. The parameter which relates PW to $SSWN$ at each age will now depend on the age of the individual and on the difference between the year of observation and the year when the reform took place.

In order to keep the exposition simpler, we consider the case where the rate of interest is equal to the subjective rate of discount, and the level of consumption is constant during the individual's lifetime. In terms of the variables of equations (7.6) and (7.7), the value for C_1 is given by:

$$C_i = \frac{1-R}{1-R^T}(RV + NPW)\ i = 1, 2, \ldots, T \tag{7.8}$$

where RV is lifetime income, and $SSWN$ is net social security wealth. These variables are defined respectively as:

$$RV = \sum_{t=1}^{L} (1-\alpha)\ W_t R^{t-1}$$

and

$$SSWN = -\sum_{t=1}^{L} \alpha W_t R^{t-1} + \sum_{t=L+1}^{T} B_t R^{t-1}.$$

Given that $C_t = C_{t+i}$, $i=1, 2,.., T-1$, we can derive the equation for the accumulation of wealth at time j, when no reform takes place from time 1 to time j:

$$PW_j = \sum_{t=1}^{j}(1-\alpha)\frac{W_t}{R^{t-1}} + R\frac{1-R^{-j}}{1-R^T}(\mathrm{RV} + SSWN). \tag{7.9}$$

The equation relates the private wealth at time j to the present value of earnings received from $t=1$ to $t=j$ and to the total life-cycle wealth (the sum of lifetime income and net social security wealth), which an individual expects at the beginning of his/her working life. The derivative of PW with respect to $SSWN$ is negative and depends on the age of the individual.

Let us consider now a situation where at time k, $k<j$, a reform in the pension system takes place. How will the individual modify his/her optimal path of consumption, saving and wealth accumulation, if he/she can re-programme them at time k? Intuitively, we can imagine that the individual at time k, when the reform is announced, starts a new programme of consumption, where the lifetime resources are the inherited wealth at time k, the present value at time k of the sum of future earnings and the present value of the new net social security wealth ($SSWN$*). We can re-write the maximizaton problem at time k as:

$$\max\sum_{t=k}^{T}\gamma^{t-k}\log C_t \tag{7.10}$$

subject to:

$$\sum_{t=k}^{T} C_t R^{t-k} \leq \frac{PW_{k-1}}{R} + \sum_{t=k}^{L}(1-\alpha)W_t R^{t-k} + \sum_{t=L+1}^{T} B_t R^{t-k}. \tag{7.11}$$

Now, the optimal value of consumption C_k depends on the total wealth at time k, and on the number of years during which the individual must divide these resources. We are interested in the level of private wealth accumulation which derives from the optimal path of consumption and saving at time $j>k$. If we substitute the new re-programmed value for consumption from k to j in the private wealth accumulation equation, we get:

$$PW_j = \frac{PW_{k-1}}{R^{j-k-1}} + \sum_{t=k}^{j}\frac{(1-\alpha)W_t}{R^{j-t}} + R\frac{1-R^{k-j}}{1-R^{T-k}}\left(\frac{PW_{k-1}}{R} + RV + SSWN^*\right) \tag{7.12}$$

where $SSWN$* is the value of the net social security wealth after the reform. The first two terms on the right-hand side of the equation represent the amount of resources which the individual would have accumulated until time j, had the consumption been equal to zero. The last term on the right-hand side measures the present value of consumption realised until j. We use the term $R(1-R^{k-j}/1-R^{T-k})$ as the adjustment factor in the regression of private wealth against net social security wealth; k is the age of the individual at the time of the reform, and j is the age of the individual at the year of the observation j.[19]

Estimation of the private to social security wealth substitution

On the basis of the variable that we have just described we shall estimate the relationship between net social security wealth and private wealth, on the survey data covering the 1991–98 period. The basic regression has the following structure:

$$PW/Y^P = \alpha\, \boldsymbol{X} + \beta\ SSWN/Y^P + u,$$

where W/Y^P, the dependent variable, is the ratio of household real and financial wealth net of debts to the permanent income of the household, Y^P; $\boldsymbol{X}$ is a vector of demographic control variables, and $SSWN$ is net social security wealth of the household. The construction of household permanent income has been described above.

To isolate the effect of social security wealth on the other components of household wealth, the independent variables in vector $\boldsymbol{X}$ should be associated with the evolution of wealth over the life cycle: we have included a third-order polynomial in the age of the head of the household, two dummies for his/her education level (primary, tertiary), two dummies for the area of residence (central or southern Italy), family size, the number of income recipients, and year dummies.

$SSWN$ is the net social security wealth of the household, obtained by summing the levels of social security wealth of the head and of the spouse, if present. Social security wealth is, as already explained, net of the present value of future social security contributions. The value of social security wealth has been adjusted taking into account the Attanasio-Gale suggestions, which we referred to above.

The sample includes those households whose head is aged between 21 and 60, and receives a positive amount of labour income. Moreover, we consider only households with no more than two earners. Families with an exceedingly high income or wealth are also excluded.[20] Two different estimation methods are used: OLS (with robust standard errors) and quantile regression, allowing the degree of offset between the two forms of wealth at different levels of the wealth-income distribution to be estimated.[21]

Table 7.21 provides a broad view of the substitution coefficients (the βs in the previous expression) for various subgroups of the population; the first row refers to the whole sample, then the coefficients are shown for households classified according to the occupational condition of the head. All coefficients are negative and significant; we therefore find support for the hypothesis of a negative relationship between pension and private

Table 7.21 Substitution coefficient between net social security wealth and private wealth

	With Gale correction		Without Gale correction	
	OLS	Quantile regression (50%)	OLS	Quantile regression (50%)
a. All households (b + e)	−0.6045	−0.4691	−0.3183	−0.2568
	(0.061)	(0.054)	(0.039)	(0.033)
b. Employees (c + d)	−0.5368	−0.4384	−0.3123	−0.2627
	(0.054)	(0.054)	(0.035)	(0.031)
c. Private sector employees	−0.4985	−0.4191	−0.2913	−0.2367
	(0.078)	(0.062)	(0.043)	(0.035)
d. Public sector employees	−0.6018	−0.4847	−0.3512	−0.3146
	(0.090)	(0.062)	(0.057)	(0.047)
e. Self-employed	−0.9215	−0.8752	−0.6656	−0.6607
	(0.192)	(0.144)	(0.115)	(0.081)

Note: Standard errors in parentheses.

wealth, although the degree of substitution is lower than 100 per cent. For the whole sample, the coefficient is greater than the values found in previous studies on Italian data (Brugiavini 1987, Jappelli 1995). The use of quantile regression, which should reduce the impact of the exceedingly wealthy, tends to lower the absolute value of the estimated coefficients, while the application of the correction proposed by Gale (1998) and Attanasio and Rohwedder (2001) has, as expected, the effect of raising the identified degree of offset.

Results for the complete regressions (with SSW/Y^P always multiplied by the Attanasio-Gale correction) on the whole sample are illustrated in the Table 7.22, which shows the coefficients estimated first by OLS, and then by quantile regressions, calculated at the 25th percentile, the median and the 75th percentile of the distribution of the ratio W/Y^P.

As for the OLS equations, the overall degree of substitution is estimated at –0.56, very significantly different from zero. Quantile regressions show that richer households are more responsive to changes in social security wealth: the offset goes from –0.25 for households at the 25th percentile, to –0.47 for those at the median, to –0.64 for households in the 75th percentile of the wealth–income distribution.

Besides this basic specification, we have carried out the same regression

Table 7.22 Relationship between social security wealth and private wealth, all households

	OLS		Quantile regr. (0.25)		Quantile regr. (0.5)		Quantile regr. (0.75)	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
SSW/Y^P	−0.6045	0.0615	−0.3566	0.0270	−0.4691	0.0536	−0.7561	0.0615
Age	0.2358	0.0170	0.1483	0.0064	0.1945	0.0142	0.2860	0.0165
Age^2	0.0038	0.0014	0.0016	0.0006	0.0005	0.0013	0.0044	0.0015
Age^3	0.00002	0.00004	−0.00004	0.00002	−0.00003	0.00005	0.00001	0.00005
Primary	−0.4658	0.0751	−0.4202	0.0392	−0.6311	0.0815	−0.5549	0.0938
Tertiary	0.0590	0.1023	0.2377	0.0573	0.1708	0.1158	0.2180	0.1372
Centre	0.0588	0.0893	−0.0307	0.0470	0.1568	0.0952	0.2234	0.1090
South	−0.0174	0.0864	−0.0351	0.0413	−0.0433	0.0885	−0.1053	0.1036
N members	0.0827	0.0325	0.0742	0.0167	0.0874	0.0335	0.1138	0.0380
N earners	−0.3910	0.0735	0.2921	0.0375	−0.0560	0.0770	−0.7358	0.0895
1991	0.2310	0.1185	0.2147	0.0569	0.5472	0.1185	0.2003	0.1348
1993	0.1826	0.1206	0.1655	0.0602	0.6597	0.1255	0.2272	0.1465
1995	−0.2834	0.1208	−0.0164	0.0593	0.2260	0.1235	−0.4664	0.1441
1998	−0.6385	0.1274	−0.1074	0.0623	−0.1173	0.1300	−0.9641	0.1541
2000	−0.9973	0.1259	−0.2615	0.0639	−0.4920	0.1327	−1.3950	0.1534
Cons	6.4595	0.2811	1.9188	0.1223	4.8224	0.2493	9.0049	0.2861
R^2	0.0930		0.047		0.065		0.07	

Note: Age is defined as age of the head: 50.

on subsamples defined according to certain characteristics of interest of the households. Table 7.23 shows, for example, that the OLS estimation on the subsample of households with a head aged between 51 and 60 provides a coefficient of the degree of substitution between *SSW* and private wealth of –0.6259, which becomes –0.5159 with median regression. Almost all coefficients are negative, as expected, and above –1. This first set of results shows that the offset is negative only for households headed by persons aged 40 years or more, while younger households actually display a positive coefficient. This result is consistent with those obtained earlier, in particular with the findings of negative cohort effects in the saving rate for younger generations, and of a limited reaction in the saving rate of the young after pension reform. Households headed by workers aged more

Table 7.23 Substitution coefficient between net social security wealth and non-social security wealth for different groups of the whole household sample

	OLS		Quantile regression	
	Coeff.	s.e.	Coeff.	s.e.
Whole sample	−0.6045	0.061	−0.4691	0.054
Age of head				
21–40	1.1696	0.134	1.2136	0.122
41–50	−0.2184	0.077	−0.2526	0.078
51–60	−0.5888	0.077	−0.5039	0.067
Education of head				
Primary	−0.5699	0.085	−0.4130	0.064
Secondary	−0.7210	0.100	−0.6100	0.072
Tertiary	−0.596	0.164	−0.6058	0.120
Computation rule of pension of head				
Earnings related	−0.8510	0.069	−0.7674	0.045
Contribution related	0.5576	0.166	0.8399	0.152
Year				
1989	−0.7960	0.162	−0.8741	0.095
1991	−1.0002	0.151	−0.868	0.113
1993	−0.5784	0.145	−0.4629	0.119
1995	−0.5479	0.149	−0.3423	0.129
1998	−0.4382	0.132	−0.4440	0.096
2000	−0.2864	0.157	−0.0743	0.130

Note: These coefficients have been estimated from separate regressions, each containing the same set of control variables as shown in Table 7.22. The only exception is that in the regressions for different age groups the age polynomial has not been included.

than 40 years are therefore more reactive to changes in social security wealth. The offset, however, is below 100 per cent. Younger households could suffer from the presence of liquidity constraints, which make it difficult to change the accumulation pattern of financial or real wealth. Regressions carried out separately on samples with different education levels do not produce significant behavioural differences; only in the quantile regression is the coefficient of the more educated greater than that of households with only primary education.

The third set of regressions in Table 7.24 checks whether there is a significant difference in the reaction of households affected differently by the pension reforms; in particular, the sample has been split between those households whose head has retained the old pension computation method (earnings-related defined benefit) and the households whose head has had to shift to the new contributory regime (notional defined contribution). The first group is composed of heads with a contributory record of at least 16 years in 1993 (18 in 1995, 21 in 1998). Only for this first group, that is, the older part of the labour force, is the offset coefficient negative and significant, in accordance with results differentiated by age of the head. Finally, estimations on the separate cross-sections of the four years show a tendency towards the reduction of the coefficient in more recent years. The result that older households turn out to be more reactive than younger ones is contrary to what one could a priori expect, because of the greater reduction in social security wealth for younger generations, but is consistent with the evidence obtained from the difference-in-differences estimations shown above.

Results from the subsample of employees are shown in Table 7.25. The overall substitution coefficient is negative and significant, and increasing in absolute value in the three quantile regressions as before. In general, all previous patterns of the coefficients are confirmed, which should not be surprising, given that employees represent 80 per cent of the dimension of total sample. Interestingly, the offset coefficient declines in absolute terms moving from 1991 to 1998.

Conclusions

In the nineties, Italy experienced a U-turn in economic policy management, and in the structural conditions of the markets:

- liberalisation of the capital markets;
- strong reduction in the government budget deficit;
- three steps reform of the public pay-go pension system;
- privatisation of almost all state-owned firms;

Table 7.24 Relationship between social security wealth and private wealth: employee households

	OLS		Quantile regr. (0.25)		Quantile regr. (0.5)		Quantile regr. (0.75)	
	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.	Coeff.	s.e.
SSW/YP	−0.5369	0.0591	−0.2406	0.0325	−0.4384	0.0535	−0.7324	0.0673
Age	0.1832	0.0168	0.0843	0.0084	0.1464	0.0144	0.2399	0.0185
Age^2	−0.0003	0.0012	−0.0011	0.0007	−0.0042	0.0013	0.0008	0.0016
Age^3	−0.00008	0.00004	−0.00008	0.00003	−0.00017	0.00005	−0.00007	0.00006
Primary	−0.6986	0.0680	−0.4445	0.0460	−0.7464	0.0775	−0.8024	0.0987
Tertiary	0.0909	0.0987	0.3027	0.0658	0.0089	0.1087	0.1976	0.1402
Centre	0.1603	0.0821	−0.0142	0.0555	0.2042	0.0908	0.4848	0.1143
South	−0.0752	0.0790	−0.0584	0.0494	−0.0369	0.0847	0.1175	0.1096
N members	0.0845	0.0306	0.0719	0.0200	0.1039	0.0326	0.1099	0.0400
N earners	−0.0321	0.0680	0.3683	0.0441	0.2866	0.0733	−0.3485	0.0940
1991	0.4994	0.1049	0.2669	0.0686	0.5436	0.1155	0.4890	0.1446
1993	0.5191	0.1099	0.2607	0.0720	0.6994	0.1212	0.5654	0.1529
1995	0.0625	0.1085	0.0273	0.0707	0.3104	0.1185	−0.0561	0.1518
1998	−0.2039	0.1207	0.0131	0.0728	0.0622	0.1241	−0.4992	0.1621
2000	−0.6317	0.1125	−0.1011	0.0766	−0.4024	0.1280	−0.8980	0.1628
Cons.	5.1120	0.2665	1.1848	0.1449	3.8350	0.2431	7.4768	0.3114
R^2	0.108		0.049		0.078		0.076	

Note: Age is defined as age of the head: 50.

Table 7.25 Substitution coefficient between net social security wealth and private wealth for different groups of employees

	OLS		Quantile regression	
	Coeff.	s.e.	Coeff.	s.e.
All depend. work.	−0.5368	0.059	−0.4384	0.054
Age of head				
31–40	0.8717	0.135	1.1082	0.119
41–50	−0.1891	0.078	−0.2380	0.064
51–60	−0.4911	0.071	−0.4710	0.067
Education of head				
Primary	−0.4140	0.082	−0.3522	0.073
Secondary	−0.7920	0.096	−0.6380	0.066
Tertiary	−0.5550	0.164	−0.0526	0.160
Head subject or not to pension reforms				
Not subject	−0.7630	0.065	−0.7264	0.054
Subject	0.3750	0.165	0.7050	0.170
Year				
1989	−1.1860	0.154	−1.1398	0.099
1991	−0.8509	0.138	−0.8839	0.068
1993	−0.5140	0.157	−0.3416	0.117
1995	−0.3720	0.149	−0.2167	0.119
1998	−0.1648	0.141	−0.2332	0.109
2000	−0.1532	0.110	−0.0334	0.171

Note: These coefficients have been estimated from separate regressions, each containing the same set of control variables as shown above. The only exception is that in the regressions for different age groups the age polynomial has not been included.

- liberalisation of some of the markets for public utilities; and
- more flexibility in the labour market legislation.

The main features of the pension system reform were:

- retirees no longer share technical progress with current working generations, that is, benefits are no longer indexed to real wage growth;
- benefits will be computed according to a notional DC system, which is fairer from an actuarial point of view;
- once the new system is fully operative, all workers (private and public employees, and the self-employed) get the same return on contributions paid;
- under the new regime, the legal age for retirement is flexible, and the

replacement rate will diminish to a larger extent for workers retiring earlier; and
- early retirement without actuarial adjustment will be abolished, during the current and subsequent decades.

The transition to the new pension regime will be a lengthy process, hence the working population will be affected to different degrees by the reform, according both to the seniority of each worker, and to the sector he/she belongs to.

As far as *aggregate* propensity to save is concerned, households do not seem to have reacted significantly to the U-turn. At first sight, the propensity to save out of household disposable income has continued declining along the general trend of the 1980s. Different definitions of income (either net of interest revenue, or gross of capital gain) allow either a stabilisation, or even an increase in the propensity to save to emerge, during the second half of the 1990s.

On macroeconomic grounds, three main forces might have been at work influencing the behaviour of the propensity to save to a different extent, and in different directions, in recent years:

- the structural impact of the budget consolidation, and the consequent reduction in the real return on government bonds denominated in domestic currency;
- the impact of the reduction in social security wealth, as a consequence of the reform of the pay-go system; and
- the long-wave change in the intertemporal preferences of the population, as new cohorts enter their working-age period with greater impatience *vis-à-vis* the future.

This research project was intended to investigate whether, below the macroeconomic surface, the microeconomic behaviour of Italian households reveals possible reactions of the individual propensity to save and of the accumulation of private wealth, to the pension system reforms.

The data of the Bank of Italy surveys that we have used in this research have revealed interesting features of Italian household behaviour. We shall start with the *age profile of the propensity to save*:

- it has a peak around the age of 60 years, in the neighbourhood of 25 per cent, and declines only mildly at older ages (not more than 2/2.5 percentage points);
- as new cohorts age, they reduce their propensity to save with respect to the older cohorts by about one percentage point on average, when they reach the age bracket of 40–60 years; and

- no systematic drift in the profile emerges from the estimate, independently of the cohort effect.

As far as the *age profile of wealth to income ratio* is concerned,

- financial wealth reaches a peak around the age of 65, but does not decline to a relevant extent as the household head gets older; and
- the age profile of financial wealth has definitely drifted upward, as a result of general macroeconomic effects; this happened to such an extent that if we try to identify a separate cohort effect, it assumes an incredibly high value.

Even stronger positive cohort and time effects emerge for the holding of risky assets. These results give some clues about the role of *inter vivos* transfers: younger cohorts seem able to have both a lower propensity to save, and a higher wealth to income ratio, as they started with a higher real and financial resource endowment than the older cohorts. The implications of these results for the future aggregate propensity to save seem to suggest that the behavioural components of the new cohorts might be more important than the change in the demographic composition of the population.

Table 7.26 shows that the pure change in the age composition of the population implies a stability of the aggregate saving rate for the whole period; this stability is due to two contrasting effects: on the one hand, the increase in the average age of the working population, which would imply an increase in the aggregate saving rate; on the other, the increasing share of the dependent elderly tends to reduce it.

Table 7.26 also shows that if we consider the estimated negative cohort effect on saving, the ageing of the already born cohorts implies a reduction in the aggregate propensity to save between 2.5 and 5 percentage points between 2000 and 2050, depending on whether we consider the cohort effects estimated with or without demographic controls.

According to a simple Solow growth accounting model, given the current features of the Italian economy,[22] and the expected change both in the size of the population and in the share of the working population, the steady-state equilibrium propensity to save could decrease by about 2.5 percentage points in the next 40 years. This means that it is the behavioural component of the demographic effect, rather than the size of the pure composition effect, which might create possible problems in future decades. This partial conclusion suggests that our analysis of the behaviour of household saving, as a reaction to the public pension system reform and to the new offer of occupational pension funds should be extended.

Table 7.26 Aggregate saving rate as a result of the expected change in the age composition of the population

	2000	2005	2010	2015	2020	2025	2030	2035	2040	2045	2050
Without cohort effects	0.278	0.278	0.279	0.280	0.282	0.284	0.285	0.284	0.283	0.282	0.281
With cohort effects											
Cohort effects estimated without demographic controls	0.278	0.274	0.270	0.271	0.270	0.269	0.267	0.263	0.261	0.257	0.252
Cohort effects estimated with demographic controls	0.278	0.277	0.275	0.271	0.268	0.262	0.256	0.251	0.246	0.239	0.226

The long transition to the new pension system implies that different groups of workers will be affected to a different extent by the reforms. When we compare the behaviour of the household groups that we assume to have experienced the reform with the control group that we assume has not, the *difference-in-differences* analysis suggests:

- when the first toughest step of the reform was enacted in 1992, the groups affected by the change seem to have reacted with an increase of about six percentage points in the relative saving rate;[23]
- eight years after the reform, in 2000, no relevant differential change in propensity to save for the groups presumably affected by the reform could be detected; which seems to suggest that it is difficult to identify a specific pension reform effect within the more general effect of the government budget crisis at that time (1992–93);
- however, within the groups affected by the reform, households, whose head has had a high level of education show the highest and most persistent relative increase in the propensity to save (of course, education level is not independently distributed with respect to income level).

Since these did not seem to be exhaustive answers, we proceeded to compute individual social security wealth, and estimate substitution coefficients of private to social security wealth.

The computation of the present value of individual social security wealth net of the present value of future contributions shows that:

- on the whole, the aggregate liabilities of the social security system *vis-à-vis* the working labour force halved between 1991, when it was twice the GDP, and 2000;
- on an individual basis, employees experienced the strongest reduction as a consequence of the 1992 reform;
- the three steps of the reform reduced the present value of benefit for the employees by about 19 per cent, and increased the value of contribution by 27 per cent;
- the self-employed experienced the toughest reduction with the second step of reform (the transition to a notional defined contribution);
- during the 1991–2000 period, the self-employed value of the benefits was reduced by 30 per cent, and the value of contributions was increased by 57 per cent;
- the change in net social security wealth has reduced its redistributive impact on total wealth (the ratio of the mean to the median net social

security wealth increased from 1.2 to 1.7 during the 1989–2000 period; in the same period the same ratio for private wealth decreased from 1.7 to 1.6);

- for a worker who was 16 years old in 1991, the ratio of his/her annual benefit expected at retirement to his/her average annual income during the working life was reduced by about a third, for an employee, and by more than 50 per cent, for a self-employed person, after seven years, in 1998;
- if he/she was 46 years old, the reduction was by 20 and 15 per cent, respectively;
- finally, while in 1991 the individual net social security wealth, on average, was slightly higher than the total private wealth (real plus financial wealth), in 2000 it was slightly less than 50 per cent half of the private wealth.

In other words, households have seen a significant move in the composition of their portfolio of lifetime assets from government liabilities to private sector liabilities – the more so, the younger the household.

These remarks do not allow us to understand how much of the change in the composition of the overall wealth of the household was the result of an individual voluntary change, and how much it was the result of the compulsory modification of the pension system. The econometric analysis of the relationship between private and pension wealth shows that:

- a reduction of €1 in the net present value of the pension wealth seems to induce, on average, a compensating increase in private wealth of between 32 and 60 cents;
- the richer the worker, the higher the substitution coefficient;
- the older the worker, the higher the substitution coefficient, even when corrected for the Attanasio-Gale effect. This seems to be the result of a complex interaction of factors: the older worker is on average richer and probably has a higher propensity for precautionary saving;
- the stronger the impact of the reform, the higher the substitution: the self-employed and, among employees, the civil servants, react more, using private wealth;
- the more educated the worker, the higher the substitution rate;
- the previous relationship is mainly relevant moving from primary education to secondary and tertiary, which might suggest that it is not just the mirror image of the relationship of the substitution rate to the level of private wealth; there seems to be a pre-determined level of education at which one becomes fully aware of the impact of the reforms, and of the alternatives available.

As a whole, the young do not seem to have reacted to the marked decrease in their net social security wealth; they start their working life better endowed than their predecessors, and seem to accumulate private wealth to a much larger extent than their predecessors did in the past. On the other hand, the more mature workers reacted to the pension reform, even if their net pension wealth did not change much.

The behaviour of the young, revealed by the non-existence of any offset effect of private to pension wealth, confirms the relevant negative cohort effect on saving extracted from the data. In contrast, the very cautious behaviour of the more mature generations might increase their propensity to save with ageing more than we would expect from the estimated age profile. This looks like precautionary behaviour induced by the fact that the reforms as such also spread uncertainty among groups less affected by the reforms. Having said that, we should go back to the question we raised earlier: the behavioural component of the propensity to save, as cohorts age, implies a decrease in the aggregate propensity to save in the next 40 years by about 3–4 percentage points in excess of the reduction in the equilibrium saving rate.

As such, the absence of any reaction by the young to the pension reform confirms the plausibility of the estimated cohort effects. On the other hand, there might be a further positive effect of a stronger than expected precautionary saving by the more mature working generations. This seems to suggest that if pension reforms are also aimed at increasing the propensity to save, they should not be targeted too far into the future: middle-aged workers must also be involved.

The reduction in the propensity to save of the younger cohorts, and the absence of a relevant reaction to the reduction in their net social security wealth, bring about the question of how to motivate voluntary saving in order to build a private funded pillar for the general pension system. We shall address this issue in Section 3.

3. THE ROLE OF TAX INCENTIVES IN VOLUNTARY PENSION SCHEMES IN ITALY: WHAT CAN OTHER COUNTRIES LEARN FROM THIS?[24]

This section looks at a widely debated aspect of Italian fiscal policy: the role of tax incentives in voluntary pension schemes.

The development of funded pension schemes, designed to accompany or replace the existing pay-as-you-go (pay-go) system, is on the agenda of most developed countries. In Italy, tax incentives are frequently advocated to encourage this reform. However, it is difficult to find sound theoretical and empirical support for the proposal of tax incentives from the point of view of efficiency or equity. For this reason, our description and analysis of the Italian case is preceded by a section listing the economic and equity reasons for providing tax incentives to pension saving schemes, together with a short survey of empirical evidence for the effectiveness of this instrument in the most important industrial countries.

A subsequent section then describes the evolution of voluntary pension schemes within the context of reform of the Italian pension system, including an evaluation of the tax incentives provided for by Italian law. One of the most important conclusions to emerge from our analysis is that, contrary to accepted opinion, tax incentives are only one, and not always the most appropriate, measure to be adopted by a policy designed to provide protection from the risks associated with old age.

A final section draws some policy conclusions and provides an insight into the possible lessons that a country like Japan can learn from the Italian experience.

Why Do Private Pension Plans Enjoy Preferential Tax Treatment?

The aim of pension systems in the more industrialised countries is not simply to ensure a universal minimum standard of living for the elderly. They are also designed to guarantee retired persons a living standard similar to the one they enjoyed during their working lives. Most of the recent debate about pension reforms has focused on the extent to which benefits that exceed the required minimum should be provided via public pay-go schemes or privately funded pension schemes.

Public pay-go schemes are mandatory, whereas private ones may be either mandatory or voluntary. When faced with the choice of whether to subscribe to a pension plan, people may act in a short-sighted way, and fail to save enough for their retirement needs. The pension system as a whole (both public and private) may consequently fail to maintain incomes after retirement, and so mandatory saving may be necessary.

One alternative is to provide incentives to retirement savings. The most common incentive is the preferential tax treatment of pension plans compared to other forms of saving. This incentive should encourage individuals to increase their retirement savings in a way that is in the 'public' interest. The cost to the public budget of such tax expenditure is justifiable precisely because it finances behaviour that corresponds to a public interest.

In the pension systems of Continental Europe, a central role is still played by public pay-go schemes characterised by a high replacement rate, so that pension benefits represent a high percentage of average earnings during one's working life. In the United States and the United Kingdom, on the contrary, pay-go systems guarantee a minimum income. Private pension plans, strongly supported by tax incentives, supplement this income by offering a pension that is more closely tailored to the living standards of workers. The introduction of tax-privileged pension plans or the reinforcement of existing ones is under investigation in many European countries, primarily as a consequence of the financial difficulties encountered by pay-go systems, as a result of the ageing of the population as a whole.

Characteristics of pension saving schemes

The purpose of providing tax incentives to pension plans is not that of encouraging saving *per se,* but of helping to establish the efficient size of pension savings. The exact definition of tax-favoured pension saving varies from one to another, but throughout the world qualified plans are accompanied by strict regulations designed to prevent tax avoidance and to enhance the efficiency of the incentive itself.

The most important characteristic is that deferred taxation has to be bound to certain specific features of the pension saving scheme, which, according to tax legislation, concern:

- the investment period: subscription to a pension plan is for a minimum period of time;
- access to pension benefits: access is only allowed when the beneficiary has reached a certain age, or after a suitable contribution period;
- the time pattern of benefit payment: annuities are generally preferred to lump sums upon retirement;
- the exclusion of assistance, such as the transfer of benefits to the spouse in case of premature death of the beneficiary during his/her working life; and
- the prohibiting or limiting of the right to advance payment of accumulated savings.

All these provisions show that the essential aim of this kind of saving is to stimulate the accumulation of a suitable amount of funds for the retirement period, through the imposition of constraints on behaviour which could be affected by myopia or moral hazard.

Are tax incentives really effective?

Whatever their form, tax incentives should provide beneficiaries of pension saving with a higher return than that offered by other similar forms of saving. Thus the incentive will only be effective if at least one of the following conditions is satisfied.

First, personal savings must be sufficiently elastic to the net rate of return. If this is so, the higher return provided by the tax-preferred pension plan can then produce a net flow of additional savings. The degree of savings elasticity is an empirical issue, since economic theory based on the traditional life-cycle hypothesis cannot provide any insight into the size or the direction of this relation. The idea itself of a general interest elasticity of saving can hardly be conceived of in a behavioural context if the usual prerequisites of complete rationality in intertemporal decision making, full information and the self-control of the individual in carrying out plans that mean forgoing short-term gratification are removed. A critical review of these arguments and of empirical research on saving interest elasticity can be found in Bernheim (2001), who shows that the results are contradictory or inconclusive.

Second, tax incentives should be capable of influencing the portfolio decisions of savers, inducing them to replace taxed with non-taxed assets. If there is perfect substitutability among different assets, then even a small spread between returns induced by a tax incentive will lead to the complete substitution of forms of saving. If tax advantages were circumscribed (for example, by the existence of a ceiling to the tax deductibility of contributions to the saving plan), then the best thing to do would be to contribute up to the ceiling and then allocate the portfolio to other unprivileged assets.

Many empirical studies have tried to measure the impact of taxation on the structure of household portfolios. However, only a few of them have looked at countries other than the United States; we shall consider research conducted in the United Kingdom, the Netherlands and Sweden. These studies broadly suggest that taxes do affect patterns of asset ownership (Poterba 2001), and that there is evidence of the possible impact of taxation on the choice of investment channel (asset selection). What emerges from all these studies, however, is the weak link between taxation and asset allocation (that is, the amount invested in each of the selected forms of saving).

It should be pointed out that all these studies are affected by two types of problem (Poterba 2001). First, there is a lack of suitable data, in particular for the high net worth households, whose behaviour is of crucial importance for this kind of empirical validation. Second, there is a problem of portfolio incompleteness: in other words, individual portfolios do not include all available assets. Consequently, portfolio choices must be modelled according to a two-phase strategy. The first phase requires us to ascertain why investors hold incomplete portfolios. Asset demand must then be modelled according to the assets that a household owns.

Empirical studies of tax-deferred pension schemes have dealt with both problems by trying to decide how much of savings are 'new' and how much are a consequence of the displacement of other forms of saving. Particular attention has been devoted to the study of individual retirement accounts (IRAs) and 401(k)s in the United States. These studies, however, are blighted by a number of problems involving data availability and the weight of strong maintained hypotheses.

One of the most controversial problems is that in order to examine the impact of tax incentives, one should be able to compare the savings choices made by two different groups of individuals: those who are eligible for, and those who are excluded from, saving schemes. The higher savings of those individuals subscribing to saving plans could only be seen as a sign of the positive effect of tax incentives on additional saving if both groups were homogeneous from all points of view (identical initial assets, working conditions, age, wage and so on). This can only be achieved if eligibility for saving schemes is exogenous. This, however, raises certain difficulties (see Bernheim, 2001; Scholtz 2001): many empirical studies into IRAs relate to the 1982–86 period, when eligibility rules had not been changed. Research into 401(k) plans usually presupposes the exogenous nature of eligibility, as the decision to start a 401(k) plan is up to the employer. But it has been contested that the eligibility is significantly correlated with the underlying saving propensity: an employers' decision to install a 401(k) plan may be a reaction to employee demand. That means that eligibility is indeed endogenous. As a result, estimates of the entity of the savings components (new and reallocated saving) differ significantly. Moreover, as Scholtz (2001) has pointed out, the same empirical results can be interpreted differently. Engen et al. (1996) examined many of the previous papers written on this subject and concluded that there was little concrete evidence that tax incentives had led to any significant increase in saving; Poterba et al. (1996) surveyed the same writings and reached the opposite conclusion.

On the whole, empirical research on saving motivation suggests that tax incentives have a positive effect, at least on the saving mix, but the importance of this effect requires further analysis.

Pension saving is not a perfect substitute for other forms of saving

Theory would suggest that the ability of a tax incentive to influence forms of saving is considerable if the alternatives offer a high degree of substitutability. This may not be the case with pension saving, which is usually conditioned by numerous constraints. In the theoretical case of the perfect financial market, where it is possible to save or dissave at will, at a given rate of interest, 'constrained' pension saving satisfies the perfect substitutability condition. In this case, there are no liquidity constraints by definition. If pension saving offers a higher net return than other forms of saving, then it would be possible, through arbitrage, to borrow funds at the market rate, subscribe to a pension plan and then repay debt with the pension benefits. However, individuals live in an imperfect world, and are subject to liquidity constraints which hinder the transfer of purchasing power over time. It follows that pension savings become less easily substituted for other forms of saving in the case of those individuals more widely exposed to liquidity constraints.

The non-substitutability of pension savings increases if saving has other functions as well as to realise the intertemporal transfer of resources over the course of time. The following examples should help to illustrate this point more clearly. Precautionary saving may be designed to cover the risk of unpredictable events (such as illness, unemployment and so on). This uncertainty may raise the preference for liquidity and reduce the interest rate elasticity of saving. Pension saving, as we have said, does not usually usually allow for lump-sum payments on accumulated savings, even in the case of unpredictable events. A pension saving scheme with tax incentives will not satisfy such needs.

Finally, the desire to bequeath can constitute an important reason for saving. Many nations' tax laws provide that, upon death of the insured person, the latter's pension benefits will not be completely transferred to the heirs, as the transfer would not be of a pension, but simply of assistance. This constitutes another case of imperfect substitutability and of the consequent minimal effectiveness of tax incentives for saving schemes. The problem could be met by policies aimed at removing existing constraints, or by increasing the tax incentive. In the first case, policy effectiveness strongly depends on the particular institutional framework; in the second, account must be taken of the cost to the state.

Tax incentives: their effectiveness and their distributive effects

Increasing tax incentives would seem the easiest option, but it is far from clear whether this would also be the most effective choice. Liquidity constraints which hinder access to pension plans differ in importance according to the age and economic condition of the individual: they are more stringent in the case of young people (with less purchasing power and more needs) than they are for old people; more so for the poor, less so for the rich. It is also likely that precautionary saving will be inversely related to the individual's economic situation. Raising tax incentives could have the unintentional effect of limiting pension saving mainly to older, richer individuals, who already display a greater propensity to saving and are more sensitive to return differentials, given that they are less affected by liquidity constraints. Such policies could not easily be justified on equity grounds. The problem becomes more complicated if one considers that the up-front deduction of contributions and tax-free accumulation in pension plans benefits people with higher marginal tax rates more than others.

Empirical evidence seems to confirm such doubts:

1. Figures show that a high percentage of individuals subscribing to individual accounts do so at the contribution ceiling. Traditionally, in the case of IRA accounts this proportion stands at about 70 per cent (Bernheim 2001). The maximum contribution is typical of less constrained savers, while non-limit contributors are very likely to be individuals facing liquidity (or other) constraints.
2. Recent empirical studies in both the United States and the United Kingdom reveal a positive relationship between the marginal tax rate (and therefore the income level) of savers and their involvement in tax-sheltered activities.

   Poterba and Samwick (1999) use tax rate changes as well as cross-sectional tax rate heterogeneity to identify the effects of taxation on asset demand in the 1983, 1989, 1992 and 1995 Surveys of Consumer Finances. They found that the probability of households owning tax-deferred accounts is a positive function of their marginal tax rate.

   Banks and Tanner (2001) find a similar correlation between tax status and take-up of tax-favoured saving schemes (their research does not look at pension plans, however, but at other schemes with tax incentives, such as personal equity plans: PEPs). They show that a higher rate taxpayer is more likely to be a PEP holder than a basic rate taxpayer by 6 percentage points. Their probit regressions for the ownership of tax-favoured assets compared to similar assets subject to tax (unit trusts and investment trusts) show that, taking into account other factors (such as wealth, age, education), marginal tax rates significantly affect asset choice.
3. Recent figures for the United States reveal a positive relationship between the pension plan subscription rate and the income and age of the taxpayer. Springstead and Wilson (2000) compared such rates in three existing US voluntary individual account-type plans: IRAs, 401(k)s and Thirft Saving Plan (TSP). They discovered that subscribers to these plans tend to be older, higher earners, as well as male, full-time workers, and either white or from a non-black ethnic minority, to a disproportionate degree compared with the population on the whole. The tax subsidisation of retirement savings in other countries is also gender biased, since women have fewer opportunities for full-time employment with pension benefits, they experience more frequent interruptions to their working lives due to care-giving responsibilities, and they earn less than men on average, thus accruing lower benefits.
4. As can be clearly seen from official reports and research, there is serious concern in both the United States and the United Kingdom over an

insufficient level of pension coverage, in particular for poorer people, within an institutional framework where private pension saving is highly encouraged by tax incentives. In the UK, where high tax incentives are offered on personal savings, and not only for pension purposes, HM Treasury (2001) reports that, according to Inland Revenue statistics published in 2000, 94 per cent of total personal sector wealth, including housing and pensions, is owned by the wealthiest half of the adult population, and 75 per cent is owned by the wealthiest quarter. This inequality has increased over time. In 1997, one in ten households had no form of savings at all (including housing, pensions and life assurance and all liquid financial savings apart from current accounts). This figure had almost doubled since the beginning of the 1980s.

Scholtz (2001) investigates the adequacy of savings for retirement purposes in the United States, using estimates of the percentage of average consumption throughout a household's working life that could be financed by transforming all pension assets into annuities upon retirement. This consumption replacement percentage varies a lot from one income decile to the next. The median percentage of households in the bottom decile reaches 46.3 per cent (which includes housing wealth). Scholtz also shows that for the typical household in the two bottom lifetime-income deciles, social security is in reality the only source of income for retirement consumption.

The above results suggest that tax incentives alone are not the most effective instrument with which to promote pension saving plans, particularly if policy target savers include poorer, younger people faced with stronger liquidity constraints.

Are there any alternatives?

Even though the effectiveness of tax incentives is somewhat limited, it does not follow that the only alternative is to reduce the voluntary content of the pension system in favour of forms of compulsory saving, as is sometimes suggested in official reports (see, for example, US Department of Labor 1998 and the UK's Secretary of State for Social Security 1998).

Among alternative answers to the problem of encouraging subscription to voluntary pension plans, we would like to give particular attention to two that seem more suitable for those in the lower-income bracket.

A first set of policy measures involves the control of the administrative cost of pension plans. As has been frequently claimed, in regard to the United States, Europe and Latin America (often presented in the 1990s as a testing ground for the privatisation of public pension systems), when the administrative costs of (individual) pension plans are correctly accounted, the gross rate of return is considerably reduced. The reduction is highly variable, and may reach 40 per cent of the gross return. Such aspects deserve careful consideration, as the administrative costs can offset, and in some cases exceed, the tax benefit.

Moreover, administrative costs depend on the personal characteristics of those subscribing to the saving plans. Poorer people, who usually contribute less and for shorter periods, are penalised to a greater extent by fixed costs (like initial costs), which are independent of the length and entity of the contribution (Cook and Johnson, 2000; HM Treasury 2001). The institution of schemes such as Stakeholder pension plans introduced in the UK in 2001, which do not penalise irregular or shorter contribution periods, could be the right answer.

A second set of policy proposals relates to the promotion of greater public information and financial education. Economic literature heavily underlines the usefulness of a full and clear explanation of the risks of short-sighted behaviour with regard to the needs of old age. It is also very important to give individuals all the information they need to choose the most suitable plan. These kinds of policies benefit less-educated, less-well-off people more, people who have more problems in obtaining and decoding economic information.

The Role of Taxation in the Promotion of a 'Second Pillar'

The evolution of policies such as those introduced during the last decade, designed to introduce a second pillar to the Italian pension system, can be seen as a useful case study in our evaluation of the effectiveness of tax incentives discussed in the previous section. All the reform proposals in this area were in fact based on the idea that the weakness of voluntary funded pensions was largely due to the insufficient tax incentive offered to this form of pension saving.

A brief history

This story began in 1993 with the introduction of the first pension fund regulations and the establishment of a tax regime for such funds. Until then, complementary forms of pensions were rare, but those that did exist were usually in the banking sector. There were no individual retirement accounts, with the exception of life insurance policies, providing a lump-sum payment or an annuity at the end of a predetermined period of contributions.

An important role was played, however, by an old severance pay scheme which continues to exist, the TFR (*Trattamento di Fine Rapporto*), instituted in the 1920s and calculated on the basis of an agreement between the trade unions and business associations within the field of private sector industrial relations. According to the laws governing this programme, employees are forced to save part of their wages (roughly one month's wages per year) and to lend it to the firm where they work. At the retirement date, or in the case of dismissal, an employee gets back the accrued contributions, capitalised at a rate of 1.5 per cent plus 0.75 per cent of the inflation rate. The TFR is funded by a book reserve.

The 1993 reforms emerged at a difficult time for Italian public finances. The measures taken to achieve fiscal consolidation also included reductions in the pension benefits of the generous public pay-go system (see Section 2). One of the policy goals was to encourage a private pension system. The central idea, which was to permeate all subsequent developments, was to replace the TFR with pension funds, mainly closed funds, promoted by firms and unions, leaving any decisions about the terms of the substitution to employers and employees (represented by the business associations and the unions, respectively). Note that in the Italian pension system, if we add the extremely high, public old-age pension contributions (32.7 per cent of income) and the TFR contribution rate (7 per cent of income), we get the highest level of contributions, and consequently of pension benefits, in Europe. Given this situation, for dependent workers, there is very little room at all for any other form of pension.[25]

The abandonment of the TFR was thought to be a good option for workers, who could reasonably expect higher returns from investing in a funded pension plan. The main resistance to such a plan was expected to come from employers, who lose a low-cost source of finance. Firms, however, could expect to take advantage of the creation of a broader financial market and the opportunity to use their contributions as fringe benefits for employees.

The most attractive aspects of the replacement of the TFR by pension funds are as follows:

- funds may be supplied to all workers within the private sector, with no exceptions, as all of them benefit from the TFR;
- this plan can generate substantial funding for the private pension system. TFR stock has been estimated at €30 billion, while the annual flow of contributions is close to €15 billion (1.3 per cent of GDP); and
- it has no significant repercussions on the public budget, as TFR already enjoys a tax-privileged treatment of the EET kind (exemption of contributions, exemption of accumulation, taxation of benefits). As the replacement of TFR should supply the main source of funding for the second pillar, tax incentives to employer and employee contributions to the new private pension scheme could not

be of a particularly generous nature: such benefits will be enjoyed in proportion to TFR substitution.

The path to reform was not an easy one, owing to the emergence of strong opposition to the abandonment of the TFR scheme. New, funded, voluntary pension schemes did not take off, and the main reason for this was the insufficient level of tax incentives.

With the major pension reform of 1995, which, as seen in Section 2, maintaining a pay-go system, radically changed the retributive system into one characterised by a strict tie between pension benefits and contributions, tax incentives to voluntarily-funded schemes were widened. The tax frame which emerged from this reform does not really correspond to any rational design: it appears a hybrid system, where the tax benefit depends on the source of contributions (employer, employee, TFR), on the investment policies of the managing fund, and on the form of benefit (lump sum or annuities). No individual saving accounts other than life insurance policies were permitted, and these maintained their previous tax-favoured regulation. The end effect was that the 1995 reform also failed to encourage private pension schemes, while demand for the more generous tax treatment of pension savings remains strong.

In the year 2000, at the end of a complex and radical reform of the whole tax system, fiscal regulation of privately-funded pension schemes was radically revised and rationalised. The cornerstone of the new system was the introduction of a *single* tax regime covering *all* forms of pension saving: open and closed pension funds, TFR, individual life insurance governed by the same constraints as other pension savings, and individual pension plans also open to subscribers other than employees, the self-employed and individual entrepreneurs. The tax law establishes the maximum amount of fiscal favour that can be granted to the beneficiary, who can opt for one of a number of different pension savings schemes.

The new system is not biased towards or against TFR or any other voluntary pension scheme. In order to enjoy the tax benefit a number of constraints must be met:

- minimum contribution period (15 years);
- pension benefits can only be enjoyed at retirement age; and
- only one-third of accumulated savings can be withdrawn in the form of capital upon retirement.

The new, single tax regime has been designed along ETT (Exempt-Exempt-Taxed) lines.

1. Contributions are tax deductible from the personal tax base within much higher ceilings than before: 12 per cent of the wage with an absolute ceiling of €5164. Once again the deduction is partially conditional upon the use of TFR as a source of finance for the funded pension scheme.
2. Capital income and capital gains matured during the accumulation phase are taxed according to the special tax regime reserved for capital income, at a constant tax rate which is slightly lower (11 per cent) than that applied to mutual funds (12.5 per cent).
3. To prevent double taxation, the tax base of benefits excludes financial income that has already been taxed during the accumulation phase. In the case of annuities, moreover, the tax base is divided into two parts: interest matured on capitalised contributions during the pension period, taxed at a proportional rate of 12.5 per cent, and a second part taxed on the basis of the marginal personal income rate. If the benefits are drawn as a lump-sum payment, the tax rate applied is the average tax rate of personal income tax for the previous five years.

Compared to the taxation of other forms of saving, the agreed tax incentive consists of the deferred taxation of contributions and the softer taxation of those returns accrued during the accumulation phase.

The tax incentives provided by the 2000 reform are, on the whole, more generous than those of the preceding tax regime. Nevertheless, the majority of financial observers considered the tax favour to be still insufficient. A tax bill has been presented for approval by parliament, designed to widen the threshold for the deduction of contributions from the personal income tax base, and thus aimed at moving towards a scheme of the EET kind.

Tax incentives in the Italian *ETT* scheme

Are present tax incentives strong enough? Will the proposals to strengthen them produce the desired effect on the evolution of pension funds? These are the main questions that remain open in the current debate.

An evaluation of the entity of tax incentives can be made if we compare it with alternative regimes. A clear, initial benchmark for Italy is the tax treatment of other forms of non-pension savings. Such tax treatment tends to be rather homogeneous, and is based on the principle of the taxation of all capital income and *matured* capital gains. There are two tax rates: 12.5 and 27 per cent: the first is applied to the majority of medium and long-term investment assets (state bonds, private securities and so on); the second is applied to short-term bonds and bank deposits. In this chapter, we shall focus on a mutual fund taxed at the 12.5 per cent rate.

Table 7.27 shows the results of the simulation of the tax burden under

Table 7.27 Internal rate of return on pension saving under alternative tax hypotheses

(a) Lump-sum capital = 0

Parameter			Italian ETT	Mutual fund	EET	No tax
		Tax rates on:				
		Contributions	0.000	0.320	0.000	0
		Capital income in accumulation period	0.110	0.125	0.000	0
		Lump-sum capital at retirement	0.256	0.000	0.256	0
		Pension components:				
		Interest in retirement period	0.125	0.125	0.320	0
		Original contributions	0.320	0.000	0.320	0
		Capital income in accumulation period	0.000	0.000	0.320	0
		Base				
Working period	32	Gross initial earning	35000	35000	35000	35000
Pension period	26	Lump-sum capital at retirement	0	0	0	0
Growth rate of earnings	2.0%	Net pension benefit	3637	2951	3573	4717
Contribution rate	4.0%					
Gross rate of interest	2.5%					
Lump-sum capital at retirement	0.0%	Internal rate of return	1.550%	0.784%	1.484%	2.500%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	38.0	68.6	40.6	0.0
		Lower rate of return				
Working period	32	Gross initial earning	35000	35000	35000	35000
Pension period	26	Lump-sum capital at retirement	0	0	0	0
Growth rate of earnings	2.0%	Net pension benefit	2284	2058	2253	3130
Contribution rate	4.0%					
Gross rate of interest	1.0%					

Lump-sum capital at retirement	0.0%	Internal rate of return	−0.153%	−0.535%	−0.203%	1.000%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	115.3	153.5	120.3	0.0
		Higher initial earning				
Working period	32	Gross initial earning	140000	140000	140000	140000
Pension period	26	Lump-sum capital at retirement	0	0	0	0
Growth rate of earnings	2.0%	Net pension benefit	13366	9546	12432	18866
Contribution rate	4.0%					
Gross rate of interest	2.5%					
Lump-sum capital at retirement	0.0%	Internal rate of return	1.239%	0.008%	0.975%	2.500%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	50.4	99.7	61.0	0.0
		Longer contribution period				
Working period	40	Gross initial earning	35000	35000	35000	35000
Pension period	18	Lump-sum capital at retirement	0	0	0	0
Growth rate of earnings	2.0%	Net pension benefit	7174	5800	7258	9306
Contribution rate	4.0%					
Gross rate of interest	2.5%					
Lump-sum capital at retirement	0.0%	Internal rate of return	1.569%	0.790%	1.611%	2.500%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	37.3	68.4	35.6	0.0
		Shorter contribution and retirement period				
Working period	20	Gross initial earning	35000	35000	35000	35000
Pension period	18	Lump-sum capital at retirement	0	0	0	0
Growth rate of earnings	2.0%	Net pension benefit	2241	1916	2199	2978
Contribution rate	4.0%					
Gross rate of interest	2.5%					
Lump-sum capital at retirement	0.0%	Internal rate of return	0.931%	0.075%	0.829%	2.500%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	62.8	97.0	66.8	0.0

Table 7.27 (*continued*)
(b) Lump-sum capital = 1/3

Parameter			Italian ETT	Investment fund	EET	No tax
		Tax rates on:				
		Contributions	0.000	0.320	0.000	0
		Capital income in accumulation period	0.110	0.125	0.000	0
		Lump-sum capital at retirement	0.256	0.000	0.256	0
		Pension components:				
		Interest in retirement period	0.125	0.125	0.320	0
		Original contributions	0.320	0.000	0.320	0
		Capital income in accumulation period	0.000	0.000	0.320	0
		Base				
Working period	32	Gross initial earning	35000	35000	35000	35000
Pension period	26	Lump-sum capital at retirement	23676	19531	25010	30233
Growth rate of earnings	2.0%	Net pension benefit	2437	1977	2394	3160
Contribution rate	4.0%					
Gross rate of interest	2.5%					
Lump sum capital at retirement	33.3%	Internal rate of return	1.428%	0.570%	1.449%	2.500%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	42.9	77.2	42.0	0.0
		Lower rate of return				
Working period	32	Gross Initial Earning	35000	35000	35000	35000
Pension period	26	Lump-sum capital at retirement	18155	15860	18556	23779
Growth rate of earnings	2.0%	Net pension benefit	1530	1379	1510	2097
Contribution rate	4.0%					
Gross rate of interest	1.0%					

Lump sum capital at retirement	33.3%	Internal rate of return	−0.284%	−0.766%	−0.295%	1.000%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	128.4	176.6	129.5	0.0
		Higher initial earning				
Working period	32	Gross initial earning	140000	140000	140000	140000
Pension period	26	Lump-sum capital at retirement	84225	63189	89563	120932
Growth rate of earnings	2.0%	Net pension benefit	8955	6396	8330	12640
Contribution rate	4.0%					
Gross rate of interest	2.5%					
Lump-sum capital at retirement	33.0%	Internal rate of return	1.031%	−0.319%	0.897%	2.500%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	58.7	112.8	64.1	0.0
		Longer contribution period				
Working period	40	Gross initial earning	35000	35000	35000	35000
Pension period	18	Lump-sum capital at retirement	35022	28818	37497	45179
Growth rate of earnings	2.0%	Net pension benefit	4807	3886	4863	6235
Contribution rate	4.0%					
Gross rate of interest	2.5%					
Lump-sum capital at retirement	33.0%	Internal rate of return	1.482%	0.643%	1.602%	2.500%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	40.7	74.3	35.9	0.0
		Shorter contribution and retirement period				
Working period	20	Gross initial earning	35000	35000	35000	35000
Pension period	18	Lump-sum capital at retirement	11588	9519	11995	14459
Growth rate of earnings	2.0%	Net pension benefit	1501	1284	1474	1995
Contribution rate	4.0%					
Gross rate of interest	2.5%					
Lump sum capital at retirement	33.0%	Internal rate of return	0.806%	−0.263%	0.795%	2.500%
(% of cumulated net benefits)		Tax wedge (% of gross rate)	67.8	110.5	68.2	0.0

the Italian tax system. The standard case considers an individual who works for 36 years and then retires for a further 26, in an inflation-free world. The individual invests a fraction of his/her salary, which grows at an annual rate of 2 per cent, in a pension scheme providing a gross rate of return of 2.5 per cent. On retirement, his/her accumulated savings can be enjoyed in the form of a constant annuity, and as a lump-sum payment (of a maximum of one-third of the total). As indicators of the tax burden, we compute the internal rate of return (IRR) from different investment plans and the tax wedge, that is, the percentage deviation of the taxed investment IRR from the corresponding tax-free IRR.

The IRRs of the Italian ETT system are then compared with alternative solutions. The first comparison (second column in the table) is with a collective mutual fund (like UICITS, Undertakings for Collective Investment in Transferable Securities; the second comparison (third column) is with a tax regime designed along EET lines. The table also shows the results of some sensibility exercises involving changes in the rate of return, the initial wage and the allocation of working/retiring time. The set of simulations shown in Table 7.26(a) refers to the case of no lump-sum payment, while those in Table 7.26(b) refer to the case of a lump-sum payment of one-third of the accumulated savings at retirement.

The results show that, for a given amount of savings from pre-tax current income, the IRR of the new saving plans is much higher than the IRR of the mutual fund. A much discussed feature of the current tax treatment of pension schemes is taxation during the accumulation phase. This is the most important departure from the EET schemes adopted in the United Kingdom, the United States and in most European countries. It may be interesting then to assess whether the Italian tax system is more burdensome than the EET model. The answer is neither easy nor unimportant. Compared with EET schemes, the Italian ETT regime taxes capital income:

- at accrual, as in the case of all other non-pension savings investments, without tax deferral (a disadvantage for the taxpayer); and
- at the same preferential tax rate, in keeping with the other substitutive tax regimes adopted for capital income, instead of the usually higher marginal tax rate of the personal income tax under the EET regime[26] (an advantage for the taxpayer).

IRRs under the Italian ETT and the EET regimes are usually very similar (see Table 7.26). The Italian ETT is generally more advantageous in cases of longer working lives and capitalisation periods.

The employee's advantage obviously grows with his/her marginal tax rate, due to the growing distance between this rate and the lower one applied to capital income.

Thus to sum up, forwarding taxation during the accumulation phase does not have any serious disadvantages. In any case, the tax wedge does not seem too large.

On the basis of our calculations, the demand for higher tax incentives for pension funds does not appear to be particularly well-grounded. In order to provide a critical assessment of the need for an increase in the tax incentives to pension savings, other aspects need to be taken into account as stated in Section 2:

1. Owing to the lack of data in Italy, there is no empirical evidence of the impact of taxation on pension fund investment choices. The sole empirical study of this subject in existence (Jappelli and Pistaferri 2001) only considers life insurance policies. It shows, however, that the reduction in the fiscal privileges granted for these investments, which mainly harmed the richer taxpayers at the beginning of the 1990s, had no significant effects on the demand for this form of saving.[27]
2. An analysis of the characteristics of subscribers to an Italian pension fund scheme shows the low participation of young people at the end of 2000. Only 10.3 per cent of them were younger than 30. This rises to 11.5 per cent in the case of open funds (COVIP, 2001). This may indicate the existence of more binding liquidity constraints for young people and insufficient financial experience.
3. No figures are available for subscribers subdivided according to income levels. On the basis of our simple simulations, we may argue, however, that the raising of the tax wedge was less severe in the tax-favoured pension regimes (from 40.6 to 61.0 in the EET case, and from 38 to 50.4 in the Italian ETT) than for the non-tax-favoured mutual fund cases (from 68.6 to 99.7). The relative advantage enjoyed by high-income subscribers is mainly due to the deduction of contributions against a higher marginal tax rate. The tax advantage is particularly high for the Italian ETT compared with the EET, thanks to the higher spread between the marginal personal income tax rate and the proportional rate applied to financial income.

An increase in the tax favour granted to pension savings could in the end subsidise those who already save, rather than new pension savers. This effect may be further enhanced by the recent government proposal whereby the improvement in the incentive takes the form of an upgrading of the ceiling on tax deductibility of contributions.

An alternative view

The data we have presented and analysed here suggest that the development of the second pillar of the pension system is not being hindered by insufficient tax benefits and so an alternative diagnosis and therapy is thus required. More effort must be made if we are to fully understand the reasons behind the hesitation of workers and employers to replace TFR with pension funds. Only after such an investigation has been completed can we then assess whether there are non-tax incentives which are more effective than the taxed ones.

It could be said that no specific measures have been devised to encourage workers to replace TFR with a pension fund. The original idea was that the potential higher returns from the pension fund would have been sufficient to induce abandonment of TFR. However, this presupposes that TFR and pension funds are perfectly substitutable, and this does not appear to be the case. As opposed to pension funds, the TFR:

- operates as an insurance policy against unemployment, which is particularly important when unemployment benefits are lacking or are inadequate as in the case of small firms;
- constitutes a form of precautionary saving. Advances of accumulated savings can be easily withdrawn in cases of important household financial decisions (buying a house, severe or prolonged health problems, a wedding and so on). This is particularly relevant with imperfect capital markets financing personal credit; and
- provides for freely disposable capital at retirement, which in the case of a standard working life can be of the order of three years' salary.

To sum up then, TFR is a form of saving that is subject to fewer constraints than pension saving. The reluctance to abandon it is thus in keeping with the findings in Section 2: tax incentives or the expectation of higher returns are not, *per se*, a sufficient reason to replace TFR with pension funds. Strengthening tax incentives may produce poor results compared with, and/or in the absence of, reforms of other aspects of the welfare system, especially the improvement of unemployment insurance.

From the company point of view, the voluntary abandonment of TFR represents the loss of a secure and cheap source of finance. In order to try and encourage companies to divert TFR into pension funds, two further steps have been taken.

First, tax rebates have been introduced in order to compensate for the potentially higher cost of market financing. The 1995 reform provided for the creation of a tax-free book reserve of 3 per cent of TFR allocated to pension funds. This measure was, however, not enough to compensate for the cost of raising funds on the market. On the other hand, a stronger incentive would have meant enormous costs for the public budget.

Second, firms were offered other financial sources as an alternative to TFR, provided they would pledge to strengthen their financial structure. This was the aim of the securitisation of TFR introduced in 1999. Listed firms, firms applying for listing or firms that accept qualified financial operators as shareholders (selling them at least 10 per cent of their capital) can transform TFR contributions into bonds and other securities, and confer them to pension funds. The proposed mechanism was not a success, however, and only managed to attract a limited group of firms.

An alternative solution has been proposed by the current centre-right government (in office from the second half of 2001), consisting of a compulsory transfer of TFR to pension funds, to be compensated by tax and contribution rebates to the firms. This proposal does not provide any compensation for the employees under the usual, albeit as yet unproven, assumption that the higher return from funds is sufficient to ease the substitution. Moreover, the use of compulsion contradicts the main underlying principle of tax incentives, that is, the encouragement of a *voluntary* choice in favour of pension saving. It is not at all clear what will come of this proposal. Many revisions and exceptions to its compulsory character have already been advanced. The abandonment of TFR appears to be a never-ending story.

What Lessons Can Be Learned From the Italian Experience?

The difficulties met by Italy in trying to transform its severance pay programme into defined-contribution (DC) pension schemes offer some interesting lessons for other countries, like Japan, who are engaged in a redesign of their private pension system. In Japan, just like in Italy, the radical reform introduced in 2001 promoted new pension plans, both defined-benefit (DB) and, for the first time, DC pension plans.

This promotion has probably been influenced by the successful US experience. But the Italian experience tells a very different story. It is interesting to briefly compare the differences between the Italian experience and that of the US 'stampede' from DB pension plans towards DC ones, which in the last 20 years has radically reshaped the private pension system in the United States. The main aspects have been well summarised by the Pension and Welfare Benefits Administration (1999). In 1975, 70.8 per cent of the 38.431 million actively involved in pension plans were subscribers to DB schemes, while the rest subscribed to DC plans. In 1985 this dropped to 46.61 per cent, and had fallen further to 35.57 per cent by 1995. The number of participants in DC plans grew from 11.217 million in 1975 to 33.244

million in 1985, and to 42.662 million in 1995. The percentage of pension assets invested by DB plans dropped from 71.5 per cent of the total in 1975 to about 50 per cent in 1995.

The reasons for this revolution have been widely analysed in the literature. An important role is deemed to have been played by the evolution of regulation, including tax treatment. Regulation of the 'terminations for reversions' had, for instance, the effect of encouraging many firms with overfunded DB schemes to divest (Ippolito 2001). The opportunity given to employees to subscribe to the new pension plans from pre-tax income played an important role here. However, the success of the DCs can mainly be accounted for by the fact that the shift from DB to DC (or to new hybrid plans, the so-called 'cash balance pensions') involved a convergence of employers' and employees' interests. Part of the change reflects the shift of employment away from large, strongly-unionised firms in the manufacturing sector, traditionally engaged in DB plans, to smaller non-unionised firms in the service sector who usually could not afford the too severely regulated DB.

Other major changes in the employment market leading to greater labour mobility, increased interest in pensions plans like DC, which grant portability upon vesting. These changes have led to changes in human resources management, reducing the role played by DB with back-loaded benefits in keeping highly skilled workers in companies. Workers appreciate the fact that benefits are more transparent under DC plans, and that the benefits are paid as a lump sum. Employers, on the other hand, do not have to bear the risk of market fluctuations.

The effectiveness of tax incentives granted to employee contributions is enhanced by the fact that employers may match employee contributions up to a specific threshold. They have a clear interest in doing so: the share of DC schemes that may be held in company stock is much higher than that allowed in the case of DB plans.

The Italian experience, in contrast, illustrates the difficulty of shifting from one pension system to another possibly more desirable one, at least from the point of view of financial market institutions, when the relative advantages enjoyed by employers and employees are neither clearly defined nor politically agreed upon.

The lessons that countries like Japan can draw from the Italian and US experiences are nonetheless similar: a successful reform strongly depends on a reasonable settlement of the interests of both employers and employees in an evolutionary institutional context. It is likely that in Japan, as in the United States, DC pension schemes will appear attractive to those people working for small companies, or to those who frequently change jobs. At the same time, they may be of interest to those companies who, due to the low rate of return achieved in recent years, face problems in the management of their DB plans. However, it is also likely that the shift will encounter obstacles in the form of the widespread underfunding of present DBs. The difficulty of sharing the losses resulting from an eventual termination of these plans could significantly hinder the reform. Further problems could rise if the ending of existing severance pay should become part of the current political agenda.

In this context, the introduction of adequate tax incentives should be carefully considered: as the US experience suggests, given the characteristics of the new DC plans, the deductibility of employee contributions can be effective. One should not, however, forget that incentives constitute a cost for the public budget: in Canada and the United Kingdom, pensions are the largest item of tax expenditure; in the United States, they are the second only to health insurance. Such tax expenditure is also of considerable entity compared with direct public spending. In the United Kingdom, for example, the total reported in the tax expenditure accounts for 1996–97 was over £10 billion, compared with £30 billion spent on state pensions (Cook and Johnson 2000). Costly incentives that disproportionately benefit higher earners may not be, as the Italian case shows, the most effective way to reach the assigned policy goals.

NOTES

1. Section 2 is authored by Massimo Baldini, Carlo Mazzaferro and Paolo Onofri; they thank Robert Clark and Yasushi Iwamoto for their remarks; Fabio Bagliano, Elsa Fornero, Francesco Giavazzi and Francesco Regonati for discussions on several sections. The authors accept sole responsibility for any errors.
2. Actually, the propensity to save shown in Figure 7.3 is $(1 - c)$, where $c = (Y_d - C)/Y_d$; C is household consumption (*NIA* [National Income Account]), and Y_d is household disposable income. Y_d computed net of interest revenues (I_r) implies $c^* = [(Y_d - I_r) - C]/(Y_d - I_r)$. Interest payments were the 12.9 per cent of household disposable income in 1993; in 2000 they were 5.8 per cent. The two times series are Prometeia elaborations on different sets of NIAs.
3. The current state of the analysis of the behaviour of the Italian propensity to save can be found in Zollino (2001).
4. Brandolini and Cannari (1994) describe the main features of the survey.
5. The surveys of the Bank of Italy identify the 'breadwinner' as the person mainly responsible for the economic conditions of the household. Not being a permanent condition of the single individual, this is likely to induce some instability in the results.
6. Age, education level, profession and sex refer to the head of the household. Age is expressed in terms of deviations from 50. Omitted dummies refer to households living in Northern Italy, headed by a woman with a primary level of education.
7. Figure 7.4 shows only the interval 25–75, so it excludes some observations for younger and older cohorts.
8. If we omit dummies relating to different years, the shape of the profile does not change. The low absolute level of this profile, if compared with that of Figure 7.7, also depends on its derivation from a regression on the individual micro data using the ratio AF/Y as

the dependent variable, while Figure 7.7 describes the evolution of the ratio between average financial wealth and average income by cohorts and years: if households with high incomes also have higher ratios of financial wealth to income, the mean of the ratios (the dependent variable in the micro-level regressions) is lower than the ratio of the means.

9. Robert Clark reasonably argued that risk preferences might be shaped by a cohort experience such as coming of age during a depression or experiencing a period of above normal returns. Given the short and very special period that our surveys refer to (1989–98), we assume that general macroeconomic effects dominate possible cohort effects coming from the past.
10. In both the specifications that we choose (regression on age and time-period dummies; regression on age and cohort indices), the absence of interactions between age effects, on the one hand, and either time or cohort effects, on the other, is equivalent to stipulating that the age profile be the same, respectively, either for each time period or for each cohort.
11. See note 8 for an explanation of the absolute level of the line in Figure 7.10.
12. Results with robust regression, not shown, are very similar.
13. No significant change is observable for the self-employed with respect to the control group. The remarks in the text do not necessarily imply that the treatment groups actually increased their propensity to save; they might have reduced it to a lower extent than the control group.
14. However, we should not ignore the fact that 1993 was an unusual year. As already mentioned, there was a deep recession and a strongly restrictive fiscal policy.
15. However, by considering subjective expectations of the retirement age, we implicitly assume a complete knowledge of the pension law by all the individuals.
16. Self-employed pension benefit in the notional DC scheme introduced in 1995 are computed on the basis of contributions equal to 20 per cent of their gross income. The contribution rate used for the computation of employees is 33 per cent.
17. We do not compute the path of wealth for the retired generation. Therefore the figure cannot be used to infer anything about the life-cycle behaviour in consumption and saving of Italian households.
18. More realistic models could also take into account uncertainty, liquidity constraints, precautionary saving and so on. Our goal here is only to show how the reaction of saving to a change in *NPW* varies with age.
19. Equation (7.10) could be further expressed as:

$$PW_j = f[PW_{k-1}k(RV + SSWN)] + g\left(\sum_{t=k}^{j} \frac{W_t}{R^{j-t}}\right) + h(RV + SSWN^*).$$

In this case, the computation of *PW* and *SSWN* at time ($k - 1$) would be informative. However, the computation of these two variables without panel data would be just a transformation of *PW* and *SSWN*.

20. Households with a wealth / permanent income ratio greater than 30 are excluded, as well as households where the head or the spouse earns more than 200 million lire per year.
21. If not otherwise specified, quantile regressions refer to the median of the distribution of the dependent variable.
22. The capital–output ratio, the capital–consumption ratio and the depreciation rate.
23. The figure mentioned refers to the difference in the behaviour of the propensity to save between ‘treatment’ and ‘control’ groups; which means that propensity to save of the treatment groups might have declined seven percentage points less than that of the control group.
24. Section 3 is authored by Paolo Bosi and Maria Cecilia Guerra.
25. Self-employed replacement rate between pension benefit and last wage is lower (50 per cent or less for younger workers). In this case there is probably more room for the development of funded pension plans.

26. We refer to standard EET regimes. Of course one could imagine an EET scheme where the capital income component of benefits is taxed at a low tax rate.
27. One of the more relevant reasons for this result is that, notwithstanding the tax changes, this form of saving enjoyed higher tax advantages than other ones.

REFERENCES

Alessie, R., A. Kaptein and A. Lusardi (1999), 'Explaining the wealth holdings of different cohorts: productivity growth and social security', TMR Progress Report, no. 6, Tillburg.

Ameriks, J. and P. Zeldes (2000), 'How do household portfolio shares vary with age?', mimeo, Wharton School, University of Pennsylvania.

Ando, A., L. Guiso and I. Visco (eds) (1994), *Saving and the Accumulation of Wealth: Essays on Italian Household and Government Saving Behaviour*, Cambridge: Cambridge University Press.

Attanasio, O. and A. Brugiavini (1997), 'Gli effetti della riforma Amato sulla formazione del risparmio' in Banca d'Italia – CIDE, *Ricerche quantitative per la politica economica*, Rome.

Attanasio, O., and S. Rohwedder (2001), 'Pension wealth and household saving: evidence from pension reform in the UK', IFS Working Paper, no.21, September.

Banks, J. and S. Tanner (2001), 'Household portfolios in the UK', in L. Guiso, M. Haliassos and T. Jappelli (eds), *Household Portfolios*, Cambridge, MA: MIT Press, pp. 112–54.

Beltrametti, L (1995), 'On pension liabilities in Italy', *Ricerche Economiche*, **49** (4), 405–28..

Beltrametti, L. (1997), *Il debito pensionistico in Italia*, Bologna: Il Mulino.

Bernheim, B.D. (2001), 'Taxation and saving', in A. Auerbach and M. Feldstein (eds), *Handbook of Public Economics*, Vol. 3, Amsterdam: North Holland, pp. 1173–250.

Borsch-Supan, A. (2001), 'Introduction to international comparison of household savings behaviour: a study of life-cycle savings in seven countries', *Research in Economics*, **55**, 1–14.

Brandolini, A. and L. Cannari (1994), 'Methodological appendix: the Bank of Italy's Survey of Household Income and Wealth', in Ando et al. (eds), pp. 307–74.

Brugiavini, A. (1987), 'Empirical evidence on wealth accumulation and the effects of pension wealth: an application to Italian cross-section data', Financial Markets Group Discussion Paper 20, London: London School of Economics.

Brugiavini, A. and M. Padula (2001), 'Too much for retirement? Saving in Italy', *Research in Economics*, 55, 39–60.

Castellino, O. (1995), 'Redistribution between and within generations in the Italian social security system', *Ricerche Economiche*, **49** (4), 21–30.

Cook, M. and P. Johnson (2000), 'Saving for retirement. how taxes and charges affect choice', Financial Services Authority, Occasional Paper Series, no. 8, London, May.

COVIP (2001), *Relazione annuale per il 2000*, Rome: Commissione Vigilanza Fondi Pensione [Italian Authority on Pension Funds].

Deaton, A. and C. Paxson (1994), 'Saving, growth and ageing in Taiwan', in D. Wise

(ed.), *Studies in the Economics of Ageing*, Chicago: University of Chicago Press, pp. 74–125.
Engen, E., W.G. Gale and J.K. Scholtz (1996), 'The illusory effects of savings incentives', *Journal of Economic Perspectives*, Fall, **10** (4), 113–38.
Franco, D. (2001), 'Italy: a never ending pension reform', in M. Feldstein and H. Siebert (eds) (2002), *Social Security Pension Reform in Europe*, Chicago: University of Chicago Press, pp. 250–312.
Gale, W.G. (1998), 'The effects of pension on household wealth: a re-evaluation of theory and evidence', *Journal of Political Economy*, **106** (4), 706–24.
Gokhale, J., L. Kotlikoff and J. Sabelhaus (1996), 'Understanding the postwar decline in United States saving: a cohort analysis', *Brooking Papers on Economic Activities*, no. 1, 315–90.
Guiso, L. and T. Jappelli (2000), 'Household portfolios in Italy', CSEF working paper no. 40, University of Salerno.
HM Treasury (2001), 'Saving and assets for all. The modernisation of Britain's tax and benefit system', no. 8, April, www.hm-treasury.gov.uk.
Ippolito, R.A. (2001), '*Tenuous property rights: the unraveling of defined benefit contracts in the US*', paper presented at the Second Annual CERP Conference, 'Pension policy harmonization in an integrating Europe', Moncalieri, Turin, 22–23 June, mimeo.
Jappelli, T. (1995), 'Does social security wealth reduce the accumulation of private wealth? Evidence from Italian survey data', *Ricerche Economiche*, **49**, 1–31.
Jappelli, T. and F. Modigliani (1998), 'The age–saving profile and the life cycle hypothesis', CSEF Working Paper no. 9, University of Salerno.
Jappelli, T. and L. Pistaferri (2001), 'Tax incentives and the demand for life insurance: evidence from Italy', CEPR, Discussion Paper no. 2787, May, London.
Miles, D. (1999), 'Modelling the impact of demographic change upon the economy', *Economic Journal*, **109**, January, 1–36.
Ministry of the Treasury (1998), *Italy's Convergence toward EMU*, Rome, January.
Paxson, C. (1996), 'Saving and growth: evidence from micro data', *European Economic Review*, **40**, 255–88.
Pension and Welfare Benefits Administration (1999), *Private Pension Plan Bulletin*, Abstract of 1995, Form 5500, www.gov/dol/pwba/public/programs/opr/bullet 1995/e_8.htm.
Poterba, J.M. (2000), 'Demographic structure and asset returns', mimeo, December.
Poterba, J.M. (2001), 'Taxation, risk-taking, and household portfolio behavior', in A. Auerbach and M. Feldstein (eds), *Handbook of Public Economics*, Vol. 3, Amsterdam: North Holland, pp. 1109–72.
Poterba, J.M. and A. Samwick (1999), 'Taxation and household portfolio composition: evidence from the 1980s and 1990s', NBER Working Paper no. 7392, Cambridge.
Poterba, J.M., S.F. Venti and D.A. Wise (1996), 'How retirement saving programs increase saving', *Journal of Economic Perspectives*, Fall, **10** (4), 91–112.
Rostagno, M. (1996), 'Il percorso della riforma 1992–1995. Nuovi indicatori di consistenza e sostenibilità per il FPLD', in F. Padoa Schioppa Kostoris (ed.), *Pensioni e risanamento della finanza pubblica*, Bologna: Il Mulino.
Scholtz, J.K. (2001), 'Achieving retirement security or a loophole ridden tax code? Saving incentives in the U.S.', paper presented at the Second Annual CERP Conference, 'Pension policy harmonization in an integrating Europe', Moncalieri, Turin, 22–23 June, mimeo.

Secretary of State for Social Security (1998), 'A new contract for welfare: partnership in pensions', London, December.
Springstead, G.R. and T.M. Wilson (2000,) 'Participation in voluntary individual savings accounts. an analysis of IRAs, 401(k)s and the TSP', *Social Security Bulletin*, **63** (1), 69–84.
US Department of Labor (1998), *The National Summit on Retirement Savings. Final Report*, Washington, DC, 4–5 June.
Van Der Noord, P. and R. Herd (1993), *Pension Liabilities in the Seven Major Economies*, Paris: OECD.
Venti, S. F. and D.A. Wise (1993), 'The wealth of cohorts: retirement saving and the changing assets of older Americans', NBER Working Paper no. 4600, Cambridge.
Zollino, F. (2001), 'Personal saving and social security in Italy: fresh evidence from a times series analysis', Bank of Italy Discussion Paper, no. 417, August, www.bancaditalia.it.

PART III

The Impact of the Demographic Transition on the World Economy

8. Incorporating demographic change in multi-country macroeconomic models: some preliminary results*

Ralph C. Bryant and Warwick J. McKibbin

1. INTRODUCTION

Population ageing is occurring at differing paces and with differing degrees of intensity in the industrialized nations of the world. Significant ageing is already under way, for example, in Japan, Italy and Germany. Major demographic changes in the United States and Canada will begin in the second decade of the twenty-first century. With a still longer lag, the demographic trends will be manifest in developing economies as well.

These changes in the demographic structure of populations will have major economic and political consequences. Government budgets, for example, will come under severe pressure. The levels of saving and investment and the overall saving–investment balance will change significantly in many regions of the world economy. These changes will be associated with large changes in cross-border flows of financial funds and goods. Whatever the proximate impetus for the changes, major adjustments in exchange rates and balance-of-payments positions will be required. The resulting international tensions will add greatly to the complications faced by national governments as they try to cope with the pressures on their domestic fiscal budgets and appropriately revise their economic policies.

The Brookings Institution has under way a major program to study issues of population ageing. As a part of this program, the two authors have organized a project on the Global Dimensions of Demographic Change. The first product of the project was a series of workshops taking stock of what the economics profession does and does not know about the macroeconomic consequences of population ageing. We prepared an initial report in the fall of 1998 on that first phase of the project (Bryant and McKibbin, 1998). As described in detail in that initial report, a substantial and growing literature has called attention to population ageing and other demographic changes associated with it. In that literature, it has been increasingly recognized that demographic shifts may profoundly influence the world economy, directly in the countries experiencing the demographic change and indirectly through changes in global trade, capital markets and exchange rates.

Though the likely importance of the international dimensions is now widely acknowledged, it is much less widely understood that existing analytical tools are inadequate for assessing the general equilibrium and cross-border consequences of demographic change. The research reported in this chapter starts from the premise that an improvement in analytical tools is a necessary condition for improving understanding of the basic cross-border issues and the policy choices facing individual nations.

Section 2 contains an overview of the influences of demographic change on consumption, saving and wealth accumulation, together with a summary of the findings of the 1998 workshops on the changes needed to introduce demographic structure into existing multi-country macroeconomic models. Section 3 outlines the analytical approach we have been following in our recent research. Section 4 summarizes how we implement this analytical approach in two simplified multi-country models. Section 5 presents preliminary simulation results. Section 6 outlines our future plans.

2. DEMOGRAPHIC INFLUENCES ON CONSUMPTION, SAVING AND WEALTH ACCUMULATION: AN OVERVIEW

Analysis of the cross-border consequences of demographic change requires examination of many interdependencies among national economies. Such analysis therefore cannot plausibly be undertaken without use of some type of a multi-country, general-equilibrium macroeconomic framework. Explicit multi-country macroeconomic models, despite their weaknesses, are unambiguously preferable to alternatives for conjecturing about the complex behavior of economies that rely on partial-equilibrium or implicit, unsystematic methods.

As discussed in our initial report, however, existing multi-country models are inadequate to the task in several critical ways. The most notable inadequacy is the existing models' failure to incorporate explicitly the effects of demographic changes. In at least three respects, researchers working with the multi-country models must improve their model structures. In particular, an improved analysis is required of the effects of demographic changes on: (i) consumption (including possibly patterns of consumption across different goods and services), saving and wealth accumulation, with appropriate allowance for the openness of national economies; (ii) the production and supply sides of national economies, again with appropriate allowance for openness; and (iii) expenditures, transfers and revenues in government budgets.

During the inventory-taking workshops initiated at the start of the project, the largest proportion of the dialogue focused on analytical methods for studying the consequences of demographic changes for consumption, saving and wealth accumulation. The economics profession holds a variety of analytical views on these topics. Yet almost all researchers in this area agree that these areas are a priority topic for further research. In this chapter we accordingly focus on improving the treatment of demographic effects on consumption, saving and wealth accumulation.

As background for our own approach to these topics, we first summarize the status of the profession's understanding. For this purpose, we recapitulate some of the material in our December 1998 report.

At a very general level, since the work of Modigliani and Brumberg (1954, 1979) and Friedman (1957), economists have accepted in broad terms the idea that many households will wish to smooth their consumption across time. The degree of this intertemporal smoothing and the main factors driving it, however, continue to be much in dispute. One set of issues concerns the manner in and degree to which households or individuals voluntarily want to act as intertemporal smoothers. A second set pertains to whether constraints external to the household inhibit agents from acting as intertemporal smoothers. In an analytical model, the average degree of smoothing taking place in the model economy also depends sensitively on how the model builder chooses to treat households' expectations, particularly about future developments in labor income and wealth.

Explicit life-cycle approaches hypothesize that consumers save little in their early years, save most in their middle-to-late working years, and then may spend down their wealth accumulation after retirement. Some researchers interpret the empirical evidence as broadly supporting the life-cycle view, including the hypothesis that consumers are patient enough to begin saving for their retirement early in their working lifetime. Other researchers, however, read the empirical evidence as providing only weak support for the life-cycle view of saving and instead supporting the hypotheses that consumers save and accumulate wealth primarily subject to the persistence of habits or to insulate consumption against uncertainty about fluctuations in income. In these latter views, the saving-for-retirement motive is much less important than habit formation or precautionary saving for uncertainty (and hence the accumulation of 'buffer-stock' assets).[1]

Some researchers argue that the life-cycle hypothesis, as studied in the context of microeconomic panel data, appears unable to account for the most prominent observed changes in countries' saving behavior. For example, the life-cycle hypothesis does not seem to do a good job of explaining the pronounced decline in the saving ratio in the United States in the last several decades. Nor, apparently, can it explain the pronounced increases in saving ratios in several Asian countries (such as China, Indonesia, South Korea, Singapore and Thailand). For these cases, one observes large time-series changes in the aggregate saving ratio – in effect, a pronounced time trend. But the demographic changes in these countries do not seem to help in accounting for the time trends. Most of the households in the US economy seem to have cut their saving at the same time, and most of the households in the relevant Asian economies seem to have increased their saving at the same time.[2]

Carroll and Weil (1994) and Carroll et al. (2000) are partial exceptions to the view that macroeconomic analysis positing life-cycle behavior performs poorly in explaining the observed correlations between saving and growth. They argue, for example, that a model postulating a dependence of consumers' utility on comparisons of consumption to a 'habit stock' determined by past consumption can do well in explaining the observed correlation between savings and growth (including in Asian countries).

Hubbard et al. (1994) are also skeptics about the view that life-cycle models are unable to shed light on actual behavior. They use an augmented life-cycle model and show that such a model can help explain, at least for the United States, the time-series comovements of consumption and income and the high historical ratio of aggregate wealth to disposable income.

One of the complicating factors leading to differences in view about the empirical validity of the life-cycle hypothesis concerns the availability and treatment of data about government pension (social security) programs. In pay-as-you-go pension programs, a major intergenerational issue is salient. It is younger and middle-aged workers who currently pay the tax revenue into the government programs, whereas it is the elderly who currently receive the benefit payments. Moreover, the operation and fiscal balance of the pension programs is importantly influenced by regulations and provisions determining eligibility. For example, provisions setting the age for early retirement eligibility have a big effect on the actual age at which workers retire, which in turn has major effects on the pressures on government budgets (Gruber and Wise, 1998).[3] Unfortunately, the available microeconomic survey data typically do not collect information on payments received by households from pension systems. This omission is a major problem for studies of the age profile of saving behavior. It could be – as argued, for example, by Meredith (1995) – that these data omissions partially explain why the life-cycle hypothesis appears not to perform well in many microeconomic studies of saving behavior.

A further difference of view in the saving–consumption literature concerns agents that in practice may not be able to borrow and lend freely as the simplified intertemporal smoothing models presume they can. The theoretical treatment and empirical importance of such liquidity-constrained households is controversial. A number of papers provide evidence suggesting that the behavior of a sizeable minority of households cannot be adequately described by intertemporal smoothing. But there is also recognition that it may not be possible to separate the features of behavior that lead to differing degrees of willingness to smooth intertemporally (for example, the precautionary saving motive) versus the effects of external impediments such as borrowing constraints that prevent some households from intertemporal smoothing.[4]

The significance of bequests is also controversial in the specification of consumption–saving and wealth accumulation. Issues at stake include the relative importance of voluntary or strategic versus involuntary, accidental bequests. In his overview in the Clarendon Lectures, Deaton (1992, p. 217) remarks: 'it now seems that bequest motives are a good deal more important than we used to think.' Kotlikoff and Summers (1981) also argued for much greater emphasis on intergenerational transfers as a driving force for aggregate capital accumulation.[5]

The treatment of agents' expectations is still another area where analytical views are heterogeneous. Few if any analysts now prefer to use backward-looking, adaptive expectations in their research. The use of forward-looking, model-consistent expectations has become common. Yet the extreme assumptions of rational, model-consistent expectations may in some ways be almost as unsatisfactory as the opposite extreme of adaptive, backward-looking expectations. Reliable empirical evidence about the pervasiveness of forward-looking behavior by individuals, households and firms is sketchy, and behavior no doubt varies significantly across different types of agents. Although forward-looking behavior of some sort is widespread, it is still an open question what proportion of consumers (if any) makes decisions in a manner that is well captured by analyses based on model-consistent expectations.

Consumption and wealth accumulation are central to both microeconomics and macroeconomics. Yet the unresolved issues are remarkably salient – even before one gets to the complications introduced by open national economies. Some degree of consensus may exist among researchers that the simplest versions of permanent-income or life-cycle models are not sufficient by themselves as adequate descriptions of saving and consumption behavior. For example, Deaton (1992, p. 217) summarizes:

> accumulating microeconomic evidence casts increasing doubt on the life-cycle hypothesis, or at least on the insights that come from the 'stripped-down' version, that saving is largely hump-saving for retirement, that consumption is based on lifetime resources, that aggregate wealth is accountable for by life-cycle saving, and that saving responds positively to productivity and population growth.

He interprets the evidence as primarily against the low-frequency intertemporal smoothing of consumption, and perceives expected future income, especially distant future income, as having 'limited relevance'.[6]

To many analysts, it still seems natural to suppose that large changes in the demographic structure of populations can have significant effects on consumption and saving. The effects should be larger, the more pervasive are life-cycle elements in the behavior determining saving. A household manifesting the traditional life-cycle behavior will have a hump-shaped lifetime profile for its saving rate: low saving in its early years (when, for example, children are very young), high saving in middle age in anticipation of retirement, and then low or even negative saving after retirement. Population ageing due to the retirement of the baby-boom workers and lower fertility should thus, other things equal, lead to a decline in the private saving rate.[7] Developing countries that experience sharp declines in infant mortality, rising fertility, and hence a surge in the population of dependent children should likewise experience higher consumption and lower saving. A nation experiencing unusually rapid growth in the labor force as baby-boom children move into the years of working age should exhibit, other things equal, a higher private saving rate. Even if simplified life-cycle considerations are less important than precautionary saving and habit persistence and even if bequest motives importantly influence savings, some significant role for life-cycle effects may be needed in analytical models, perhaps especially those aspiring to capture cross-border spillovers.[8]

Most aggregative macroeconomic models fail to allow for changes in the demographic structure of populations. The seminal contributions to growth theory in the 1950s and 1960s abstracted from such demographic factors. In effect, those models contained no children and no elderly, with the result that a faster (steady-state) rate of growth of population caused the saving rate unambiguously to rise in response to higher requirements for investment and a higher capital stock. Tobin (1967) developed a simulation model differentiating workers from retirees (but still without children). Tobin's model also predicted that faster population growth would raise the private saving rate, because the faster growth caused the population distribution to become younger (more working, saving households relative to retirees who were older and dissaving).[9] Even as theoretical growth models matured, most empirical macro models still abstracted from shifts in the demographic composition of the population. That omission persisted despite the major emphasis in consumption theory of life-cycle considerations in the microeconomic behavior of households.[10]

Another aspect of consumption which may be important for understanding the impacts of demographic shifts is changes in the composition of consumption bundles over the lifetimes of individuals. In a single good world, it is not possible to capture the impact that changes in preferences may have on the relative prices of alternative goods and services. For example, in an ageing society it is likely that the demand for health services will rise, which will cause shifts in the relative prices of these services and a change in the allocation of resources to satisfy shifting demands. Conventional single-good macroeconometric models cannot capture this type of compositional effect. Multi-sectoral multi-country models, such as the G-cubed model of McKibbin and Wilcoxen (1999), could do so if demographics are incorporated appropriately.

Issues of aggregation are important when researchers specify the consumption–saving–wealth relationships in macroeconomic models. In particular, the aggregation issues are central for getting an adequate macro specification of the demographic influences. Macro models built up from a micro theory positing a single representative agent are not easily adapted so as to incorporate demographic changes. By definition, changes in the demographic composition of the population require analysis to acknowledge the heterogeneity of agents – at the very least heterogeneity in age.[11] The spirit of overlapping-generations (OLG) models is to grapple directly with one or more dimensions of heterogeneity across agents. Proponents of the OLG modeling tradition tend to believe that macro-model specifications without allowance for heterogeneous agents are bound to be inadequate – see, for example, Kydland and Petersen (1997).[12]

For adequate macroeconomic analysis of the effects of demographic changes on consumption and wealth accumulation, is it essential to use explicit multi-cohort OLG models? Alternatively, for many analytical purposes, could it be sufficient to adopt one or another empirical 'shortcut' to modify the existing macro models that have a more conventional life-cycle, permanent-income specification and that currently either ignore demographic composition effects or else treat them in an aggregative, representative-agent approach?

Any researcher working in this area must take a stand on the issues summarized above. The issue of whether to pursue multi-cohort OLG models or whether to pursue an analytical shortcut approach is particularly difficult. We know from the international workshops held in 1998 that there is no consensus on which of the two routes has the greater probability of short-run success. From the perspective of the profession as a whole, both approaches will continue to be followed and further refined.

The trade-off facing researchers about the two routes is primarily a matter of time horizon. Multi-cohort OLG approaches that explicitly keep track of different cohorts, their saving decisions and their wealth stocks can be more rigorous theoretically. Other things being equal, a multi-cohort OLG approach thus may appear more attractive. But other things are not equal. Multi-cohort OLG models are more difficult and demanding than models that use analytical shortcuts to get demographic effects into the consumption-saving specifications in macroeconomic models. The OLG models, moreover, are likely to take much longer to advance to the stage where the models can deliver interesting empirical conclusions. The requirements of a multi-cohort specification are of course especially demanding in a model with numerous separate national economies and national currencies.

Another disadvantage of the multi-cohort OLG specification is that it might, if calibrated only to partial-equilibrium relationships derived from micro-level evidence, deliver misleading inferences about aggregative macroeconomic relationships. As Hamid Faruqee emphasized to us in a comment made after the international workshops that launched this project, it is both the virtue and the vice of a micro-level specification for individual agents that variables such as goods prices and interest rates are taken as exogenously given. Goods prices and interest rates self-evidently cannot be modeled as exogenous at an aggregative, economy-wide level. It is thus unclear whether a disaggregated OLG model, based on partial-equilibrium relationships estimated from micro-level data, will yield correct inferences and predictions for general-equilibrium, macroeconomic behavior (in the sense of being able to replicate the moments of actual macroeconomic data).

What we have labeled as 'shortcuts' in the consumption specifications of macroeconomic models exist in several forms. Masson and Tryon (1990) followed one such route in their first effort to adapt the IMF Staff's MULTIMOD to study the consequences of population ageing.[13] Meredith (1995) followed a similar approach in studying demographic changes and saving in Japan. A different shortcut approach to incorporating the effects of demographic changes in the equations of macroeconomic models is described in Fair and Dominguez (1991).[14] The staff in the Economics Department at the OECD experimented with a shortcut incorporation of demographics in a new 'Minilink' model; see Turner et al. (1998). Bryant experimented along analogous lines at Brookings, using a two-region abridgement of the IMF staff's MULTIMOD that is a precursor of the model described later in this chapter. The OECD staff and Bryant both built on a modified consumption–saving–wealth specification embodied in the Mark III revision of the IMF staff's MULTIMOD, as described in Laxton et al. (1998) and Faruqee et al. (1997).

The shortcut approaches identified in the preceding paragraph specified a negative macroeconomic link between dependency ratios and saving rates. That relationship was justified in part on the basis of regression analysis with macroeconomic data. Numerous macroeconomic studies reported that negative correlation. Thus, at least at first blush, the age profile of savings appeared more consonant with the life-cycle hypothesis at the macro than at the micro level.

Table 8.1 replicates a survey of estimates from previous studies prepared by Guy Meredith (1995, Table 4–1).[15] It is relatively easy with macroeconomic data, as the table shows, to run regressions using an aggregate saving rate as the dependent variable that yield significant coefficients on the youth dependency ratio and the elderly dependency ratio. Some of the estimated coefficients in these studies are implausibly high, suggesting for example that a 1 percentage point increase in the elderly dependency ratio could lead to a decline in the saving ratio of as much as 1 percentage point or more. But even the macroeconomic studies estimating much smaller coefficients, for example as small as -0.15 to -0.30, imply very substantial effects on saving from changes in the demographic composition of the population.

Macroeconomic estimates such as those in Table 8.1 are contentious, and justifiably so. Skeptics assert that when one adjusts such macroeconomic regressions properly for country intercepts and other econometric problems, one finds that the demographic effects again tend to be rather small, if not to go away altogether. The adjusted effects also seem to be quite heterogeneous across countries. Most important, as stressed above, there is substantial dissonance between these macroeconomic estimates and the microeconomic evidence based on household survey data.

3. A MODIFIED APPROACH FOR INCORPORATING DEMOGRAPHICS IN CONSUMPTION AND WEALTH ACCUMULATION

Hamid Faruqee and Douglas Laxton, economists at the International Monetary Fund, have recently suggested an improved 'shortcut' approach for the incorporation of demographic factors into model specifications of consumption, saving and wealth accumulation. The approach builds on their earlier work with the Blanchard (1985)–Weil (1989)–Yaari (1965) overlapping-generations framework – as in the Faruqee et al. paper (1997) and the Mark III version of MULTIMOD – but importantly modifies the manner in which demographics are treated. Our current research is exploring this

Table 8.1 Summary of previous studies' estimates of demographic effects on saving rates

Study	Data source	Effect on saving rate of a 1 percentage point rise in demographic ratio	
		Youth dependency ratio	Elderly dependency ratio
Aggregate cross-section studies			
Modigliani (1970)	–	−0.20 (3.7)	−0.88 (3.1)
Modigliani and Sterling (1983)	–	−0.13 (1.4)	−0.51 (4.3)
Feldstein (1980)	–	−0.77 (3.9)	−1.21 (2.7)
Horioka (1986)	21 OECD countries, 1976–82 average	−0.92 (4.2)	−1.61 (4.0)
Graham (1987)	24 OECD countries, 1975 or 1970–80 average	−0.87 (2.9)	0.12 (0.3)
Koskela and Viren (1989)	23 countries, 1979–83 average	−0.73 (1.7)	−0.76 (0.8)
Horioka (1991)	14 OECD countries, 1980–88 average	−0.44 (1.7)	−1.09 (2.4)
OECD (1990)	14 OECD countries, 1980–88 average	–	−0.93 (2.4)
Time-series studies:			
Shibuya (1987)	1966–83 (Japan)	–	−0.34 (3.8)
Horioka (1991)	1956–87 (Japan)	−0.30 (5.1)	−1.13 (3.7)
Masson and Tryon (1990)	1969–87 (pooled) G7 plus small industrial countries	−1.10	
Weil (1994)	1960–85 (pooled)	−0.3	−0.5
Masson et al. (1995)	1972–93 (pooled), 21 industrial countries	−0.15 to −0.30	
Meredith (1995)	Simulation model	−0.30	

Notes: Figures in parentheses are estimated *t*-statistics. The youth dependency ratio here is defined as the ratio of the population aged 0–19 to population aged 20–64. The elderly dependency ratio is defined as the ratio of the population over 64 to population aged 20–64 (except in the OECD (1990) study where it is the ratio of population over 64 to total population). See Meredith's original table for additional footnote information.

Source: Meredith (1995).

approach and adapting it further to empirical multi-region models. We believe the approach is promising and represents a marked improvement over previous shortcut efforts. The modified approach yields effects of changing demographics on saving rates that are somewhat smaller than the effects obtained in previous macroeconomic analyses. But the effects are still present, and important enough to significantly influence macroeconomic outcomes.

Essentials of the Modified Blanchard Framework for Consumption and Saving

Blanchard's seminal theoretical paper (1985) made use of work by Yaari (1965). The model was extended by, among others, Buiter (1988) and Philippe Weil (1989). A version of the basic model has been incorporated in the IMF's MULTIMOD and in several studies by IMF economists. Variants of the model appear in numerous other studies. Zhang (1996) and Bryant and Zhang (1996a, 1996b), for example, used a version of it in their research on fiscal closure rules.

In models in which agents have finite life horizons, the analysis of retirement and elderly dependency is analytically complex. To assume in a model that individuals retire at a given age and thereafter receive zero labor income introduces a discontinuity that greatly complicates the modeling of saving, consumption and wealth accumulation. Another discontinuity results in model environment when children enter the labor force and start to earn labor income at a specific given age. With finite-lived agents, no exact or even approximate aggregate consumption function can be derived by aggregating over the behaviors of individuals who differ in age and in consumption propensities.

Blanchard showed that a simplifying assumption can expedite aggregation and modeling and thereby allow a researcher to avoid the adoption of a more complex and analytically difficult multi-cohort OLG approach. Blanchard's key assumption was that each individual, throughout life and regardless of age, faces a constant probability of death, p. The expected life of an individual is thus $1/p$. With this assumption, researchers can choose a value for p anywhere between zero and a large number. If p is put at the limiting case of zero, individuals live forever and the model yields the infinite-horizon results familiar from still simpler models; values of p in the range 0.03 to 0.01 yield model 'life expectancies' in the range of 33 to 100 years.[16]

The constant-probability-of-death assumption can be combined with an assumption, based on Yaari (1965), that the economy contains life insurance companies permitting agents to costlessly make annuities contracts contingent on their death.[17] The two assumptions together permit the derivation of an aggregate consumption function without keeping explicit track of the consumption and wealth of multiple cohorts. Aggregate consumption turns out to be a relatively simple linear function of human and non-human (financial) wealth, with the marginal propensity to consume dependent on the age-invariant probability of death and individuals' rate of time preference.[18]

Blanchard's original theoretical exposition assumed for convenience that the population is stationary and there is no growth in productivity. As Blanchard stated in a footnote and as Buiter (1988) and Weil (1989) showed in detail, however, the model can be readily adapted to cover the cases of a growing population and growth in productivity.

The original theoretical exposition included a simplified method of permitting the income and saving of individuals to decline with age. In their follow-up work, Faruqee et al. (1997) introduced a more elaborate age–earnings profile for individuals' incomes. They allowed for a time profile of an individual's labor income that corresponds fairly closely to empirical observations, showing a rise with age and experience when individuals are relatively young but then eventually declining with age as individuals approach retirement years and beyond. Aggregate total labor income in their model was distributed according to empirically derived age-specific weights. Faruqee et al. avoided a sharp discontinuity occurring at retirement age by approximating the lifetime age–earnings profile with a continuous curve of exponential terms that eventually declines to zero but does not have an explicit kink at retirement age.[19]

Still another refinement – allowance for liquidity-constrained consumers – was introduced by Faruqee et al. (1997) and subsequently adopted in MULTIMOD Mark III. As noted above, the motivation for this refinement is that capital-market imperfections prevent some agents from borrowing against their future income. The relevance of such a constraint is especially likely for younger individuals who have not yet established their creditworthiness and who have insufficient collateral in the form of financial wealth. These agents denied access to credit markets are assumed to have no choice but to consume out of their current resources. With the addition of an assumption that a fraction of consumers are constrained from borrowing, overall consumption depends on the behavior of intertemporally smoothing agents consuming out of lifetime ('permanent') income *and* agents consuming out of current income. The simplifying assumption that the fraction of liquidity-constrained consumers is constant over time can be roughly justified by supposing that as one cohort graduates out of the credit-constrained class another cohort of the same size is born into it.

McKibbin and Sachs (1991), McKibbin (1997) and McKibbin and

Wilcoxen (1999) also allow for a constant fraction of consumption that is constrained by current incomes. The earlier modeling analyses by McKibbin and colleagues, however, did not incorporate Blanchard's assumption of a constant-throughout-life probability of death.

Demographics with Endogenous Modeling of Population Growth

Although the introduction of a realistic age–earnings profile by Faruqee et al. (1997) permitted life-cycle behavior to be manifest in the aggregate consumption function, their approach was still a fairly mechanical representation of life-cycle elements. The approach was a 'top-down' method of incorporating the age–earnings profile, in the following sense. The aggregate total of labor income was distributed across different age cohorts by the fixed age-specific weights derived from empirical observations. But the aggregate total of labor income itself was independent of changes in the demographic composition of the population.

The important innovation in Faruqee (2000a, 2000b) is to abandon the top-down approach and instead to build up the total of labor income from, so to speak, the 'bottom up'. The bottom-up approach permits changes in the demographic composition of the population to influence not only the allocation of total labor income across age cohorts but also the aggregate amount of total labor income itself. This modification of the Blanchard approach thereby permits more accurate study of the effects of demographic change on consumption and wealth accumulation, and hence on all macroeconomic variables.

The modified approach treats an economy-wide birth rate and death (mortality) rate as time-varying, age-invariant exogenous variables.[20] When children are ignored in the simplest version of the model economy, youth dependency does not exist. The 'birth' rate in the simplest version then has to be interpreted as the arrival of new adults into the population. In effect, adults are born at, say, age 20 and immediately enter the labor force. The model also presumes that immigration and emigration are always zero.[21]

For the model without youth dependency, the assumptions of an adult 'birth' rate and a single mortality rate lead to a simple dynamic equation for the evolution of the overall population:

$$\frac{\dot{N}(t)}{N(t)} = n(t) = b(t) - p(t) \tag{8.1}$$

where N is the level of the (adult) population and n is the population growth rate, equal to the difference between the birth rate, b, and the age-invariant death rate, p. A dot over a variable indicates the derivative with respect to time. The number of individuals belonging to a particular cohort s at time

s (the date they are born) as a proportion of the contemporaneous population is given by:

$$N(s,s) = b(s)N(s). \tag{8.2a}$$

The number of survivors from that cohort at a subsequent time $t \geq s$ is given by:

$$N(s,t) = N(s,s)e^{-\int_s^t p(v)dv}. \tag{8.2b}$$

If equation (8.1) is integrated over time, the size of the total population at any moment in time (up to a constant of integration) is:

$$N(t) = e^{\int_{-\infty}^{t} n(v)dv} = e^{\int_{-\infty}^{t} [b(v)-p(v)]dv}. \tag{8.3a}$$

Hence in this simplified world the population evolves according to the accumulation of past changes to its growth rate, which is just the past differences between the birth rate and the common-across-all-cohorts death rate. The population can also be defined in terms of the total of all existing individuals, summed across all cohorts (indexed by s):

$$N(t) = \int_{-\infty}^{t} N(s,t)ds. \tag{8.3b}$$

To have a measurement of the age composition of the population, one can define an elderly dependency ratio as the proportion of the population above a certain threshold age level. The cohort just attaining this threshold age each period is indexed by $i(t)$. The (fixed) difference between this index $i(t)$ and the present time, t, reflects the number of adult years needed to reach elderly dependency status. For example, members of the cohort reaching the model age of 45 years at time t were born and entered the labor force 45 years ago – that is, $t - i(t) = 45$.[22] Given a value for the elderly index that is constant over time, a dependency ratio can be defined as:

$$\phi(t) = \int_{-\infty}^{i(t)} \frac{N(s,t)}{N(t)} ds, \quad 0 \leq \phi \leq 1. \tag{8.4}$$

In an economy with a constant adult birth rate, the dependency ratio would also be constant. For the case where the birth rate and death rate are time varying, the dependency ratio evolves over time according to:

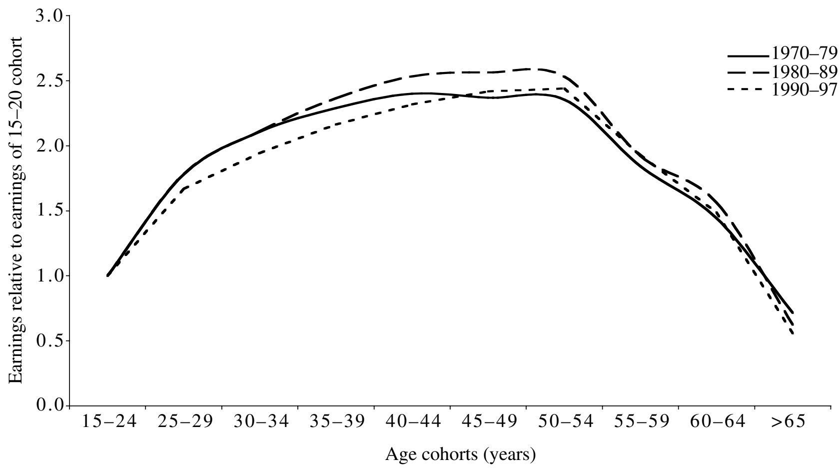
$$\dot{\phi}(t) = \frac{N[i(t),t]}{N(t)} - [p(t) + n(t)]\phi(t). \tag{8.5}$$

Intuitively, the change in the dependency ratio is determined by the relative size of new 'dependents' reaching the threshold age (the first term in (8.5)) *less* the proportion of the elderly who die in the period, $p(t)\phi(t)$, *less* a scaling term accounting for growth in the population base, $n(t)\phi(t)$.

Age–Earnings Profiles from the 'Bottom Up'

The key new component in the approach summarized here is the introduction of age-earnings profiles for labor income.

The profiles entering the theoretical specification broadly match the age–earnings profiles observed in actual datasets. After an individual enters the labor force, his or her labor income rises with age and experience, reaches a peak in late middle age, and then declines gradually for the rest of life. Figure 8.1 illustrates the typical hump shape of age–earnings profiles with Japanese data for the years 1970–97.



Note: The graph lines plot a multi-year average of earnings relative to earnings of the 15–24 age cohort (with the 15–24 cohort normalized to 1.0).

Figure 8.1 Age–earnings profiles, Japanese data 1970–1997

In real life, the labor earnings of individuals at the time of retirement fall, in some cases sharply, but often not all the way to zero. For the economy as a whole, post-retirement labor earnings of the elderly decline gradually, reaching zero for the elderly of advanced age. In the approach implemented for our model economies, it is necessary to approximate the age–earnings

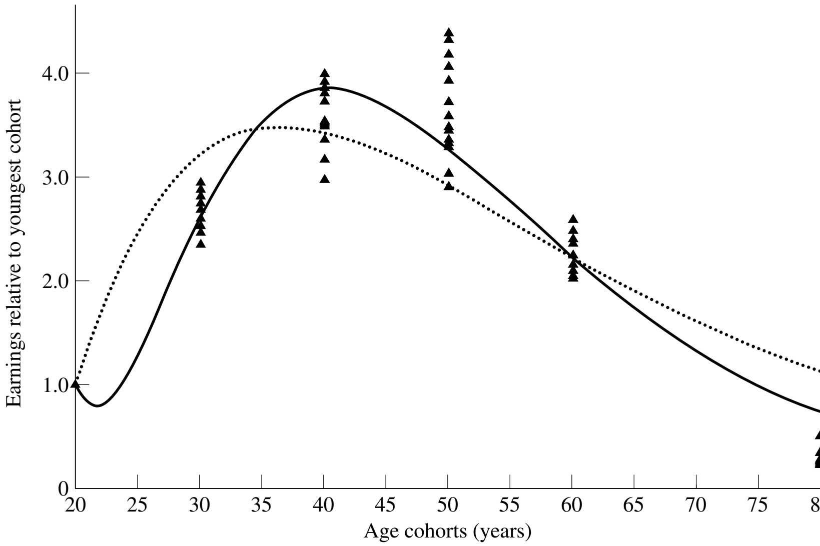


Figure 8.2 Alternative approximations of an age–earnings profile, US data 1980–1995

profile with a mathematical formulation that is continuous and permits aggregation across cohorts. Figure 8.2 shows two alternative curves that roughly approximate the shape of the humped age–earnings profiles in actual data (with the illustrative data taken from the United States).

In the model, the shape of the age–earnings profile for individuals in the economy is assumed to be the same for all individuals and unchanged through time. The labor inputs of all cohorts are inelastically supplied, though the supply varies across cohorts depending on their age. The implicit assumption is that the underlying institutional and structural features that determine variation in relative labor earnings across age groups can be taken as fairly stable in recent historical time.[23]

The composition of the population by age can change over time. Even with the assumption of an unchanged age–earnings profile, therefore, demographic shifts can generate large changes in the earnings from labor income. Total labor income is obtained by aggregating over individuals that differ in age and experience. Hence a bottom-up aggregation over individuals permits the demographic changes to influence both the aggregate level and the age distribution of labor income.

The hump-shaped profile of earnings by age influences both the supply side and the demand side of the model economies' behavior. On the supply side, the earnings profile is an indicator of the changes in a cohort's relative productivity and its supply of labor over its lifetime. On the demand side, the anticipated path of labor income determines the saving plans of consumers over their lifetimes. Hence through these life-cycle effects, changes in demographics significantly influence macroeconomic outcomes.

Specifically, the labor input of an individual cohort s at time t is assumed to be given by:

$$l(s,t) = [a_1 e^{-\alpha_1(t-s)} + a_2 e^{-\alpha_2(t-s)} + (1 - a_1 - a_2) e^{-\alpha_3(t-s)}]. \qquad (8.6)$$

The difference between the time index t and the cohort index s specifies the age of a particular cohort. The three exponential terms are a way of approximating the age–earnings profile. The parameters a_1, a_2, α_1, α_2, and α_3 are specified exogenously in modeling code. They can be estimated econometrically for individual countries, for example, by using a non-linear least squares estimation procedure with actual data for age and earnings. (Figure 8.2 shows two examples of such estimates for US data.) Loosely speaking, the first two exponential terms may be thought of as representing the decline in an individual cohort's labor supply over time as it ages and (gradually) retires. The third term can be interpreted as reflecting gains in earnings that accrue with age and experience. The restriction on the a_i terms (the third of the terms must be equal to $1 - a_1 - a_2$) embodies a normalization that the youngest cohort (for whom $s = t$) earns income equal to unity. Together the three exponential terms provide the hump-shaped profile for earnings.[24]

The earnings of an individual cohort also change over time because of general growth in labor productivity, assumed to apply uniformly to all cohorts (after adjustment for their age-specific *relative* productivities). The earnings of a particular cohort are:

$$y(s,t) = wage(t) l(s,t) \qquad (8.7)$$

where $wage(t)$ is the economy's wage rate and the wage grows through time at a constant rate of labor-augmenting technical change, μ.

If (8.6) and (8.7) are aggregated over all individual cohorts, aggregate labor income can be written as:

$$Y(t) = \int_{-\infty}^{t} wage(t) l(s,t) N(s,t) ds = wage(t) L(t) \qquad (8.8)$$

where L is aggregate labor input (adjusted for cohort-specific relative productivities). The definition of labor input for the individual cohort in (8.6) also permits one to write aggregate L as the sum of three components L_1,

L_2 and L_3 where each component reflects an exponential term in (8.6). Specifically, define:

$$L_k(t) = \int_{-\infty}^{t} l_k(s,t)N(s,t)ds \tag{8.9a}$$

where $l_k(s,t) = a_k e^{-\alpha k(t-s)}$ and $k = 1, 2, 3$ so that:

$$\dot{L}_k(t) = a_k b(t)N(t) - [\alpha_k + p(t)]L_k(t). \tag{8.9b}$$

Then one can specify a dynamic equation for aggregate labor input as:

$$\begin{aligned} \dot{L}(t) &= \dot{L}_1(t) + \dot{L}_2 + \dot{L}_3(t) \\ &= b(t)N(t) - [\alpha_1 + p(t)]L_1(t) - [\alpha_2 + p(t)]L_2(t) - [\alpha_3 + p(t)]L_3(t). \end{aligned} \tag{8.10}$$

The intuition behind (8.10) is that changes in aggregate labor input depend on the effective labor supply of new entrants to the labor force and on the death and relative productivity experiences of existing workers. Our model code uses a discretized version of equations (8.9) and (8.10) to describe the dynamic behavior of labor supply. The specific values of the five coefficients a_1, a_2, α_1, α_2 and α_3 obtained from estimating the age–earnings profile play an obviously important role in determining the movements of effective labor supply over time. They also play a critical role in the evolution of human wealth and consumption over time.

The specification of financial wealth, human wealth and consumption follows lines that are by now familiar in macroeconomic models. Individuals maximize expected utility over their lifetimes subject to an intertemporal budget constraint. Financial assets for an individual or household – $fw(s,t)$ – are determined by its saving, equal to the difference between income and consumption:

$$\dot{fw}(s,t) = \{[r(t) + p(t)]fw(s,t) + y(s,t) - \tau(s,t)\} - c(s,t). \tag{8.11}$$

In this expression for the change in financial wealth, r is the real interest rate, τ is the tax rate on labor income, $y - \tau$ is disposable labor income, and c is consumption. The term $p(t)fw(s,t)$ enters into the intertemporal budget constraint reflecting the operation of the stylized insurance-annuities market as in Yaari (1965) and Blanchard (1985).[25]

Suppose that the utility function of the individual is of the constant relative risk aversion (CRRA) form, $u[c(s,t)] = [c(s,t)^{1-\sigma} - 1]/(1-\sigma)$, where σ is the coefficient of relative risk aversion, and $1/\sigma$, is the intertemporal elasticity of substitution (IES). Solution of the consumer's utility maximization problem gives the individual's consumption as a linear function of total (financial plus human) wealth:

$$c(s,t) = \frac{1}{\Psi(t)}\,[fw(s,t) + hw(s,t)]. \tag{8.12a}$$

Here $hw(s,t)$ is a measure of the individual's human wealth (essentially the present value of future labor income); $1/\Psi(t)$ is the marginal propensity to consume out of wealth, with $\Psi(t)$ given by:

$$\Psi(t) = \int_t^{\infty} e^{\frac{1}{\sigma}\int_t^{v}\{(1-\sigma)[r(\mu)+p(\mu)]-[\theta+p(\mu)]\}d\mu}\,dv, \tag{8.12b}$$

which can be written as:

$$\dot{\Psi}(t) = -1 - \frac{1}{\sigma}\{(1-\sigma)[r(t)+p(t)] - [\theta + p(t)]\}\Psi(t). \tag{8.12c}$$

The parameter θ is the individual's rate of time preference (for simplicity assumed to be constant across all agents in the economy).

The marginal propensity to consume out of wealth in the general case of the CRRA utility function thus depends on the intertemporal elasticity of substitution and on the entire sequences of future interest rates and future adult mortality rates. This dependence is readily evident in equations (8.12b) and (8.12c). In contrast, if the utility function is of the logarithmic form where the intertemporal elasticity of substitution is forced to be unity and if the adult mortality rate is assumed to be constant rather than time varying, the marginal propensity to consume out of wealth in equation (8.12c) reduces to a constant, the sum of the time preference rate and the mortality rate ($1/\bar{\Psi} = \theta + \bar{p}$). Individual consumption then takes the simpler form:

$$c(s,t) = (\theta + \bar{p})[fw(s,t) + hw(s,t)]. \tag{8.12d}$$

For the purposes of the empirical simulations in this chapter, we report results for the simpler case where the consumption of the individual follows (8.12d) ($\sigma = 1$).

The saving behavior embodied in equation (8.11) means that individuals eventually build up financial wealth to maintain a certain level of consumption in later years. The income from accumulation of financial wealth offsets the decline in their labor income and human wealth as their labor supply falls gradually. This behavior contrasts with more traditional life-cycle models in which the elderly are presumed to run down financial assets aggressively in later life (exhibiting negative saving rates). Those traditional models, as noted earlier, are often criticized for the presumption of negative saving by the elderly, which tends to be inconsistent with the empirical facts in many countries. In the specification used here, individuals in effect target a certain level of financial wealth as a precaution against the possibility of remaining alive without sufficient labor income.

An expression for individual human wealth is given by:

$$hw(s,t) \equiv \int_t^{\infty} [y(s,z) - \tau(s,z)] e^{-\int_z^t [r(\mu) + p(\mu)] d\mu} dz. \quad (8.13)$$

A corresponding dynamic equation for the individual's human wealth can be written:

$$\dot{hw}(s,t) = [r(t) + p(t)] hw(s,t) - [y(s,t) - \tau(s,t)]. \quad (8.14)$$

In a way analogous to the exponential terms used to represent the age–earnings profile in equation (8.6), the individual's human wealth can be expressed as the sum of three components:

$$hw(s,t) = hw_1(s,t) + hw_2(s,t) + hw_3(s,t) \quad (8.15a)$$

with the components defined as:

$$hw_k(s,t) = \int_t^{\infty} [1 - \tau(t)] wage(t) l_k(s,v) e^{-\int_v^t [r(\mu) + p(\mu)] d\mu} dv \quad (8.15b)$$

for $k = 1, 2, 3$. In equation (8.15b) it is assumed that labor income taxes are proportional to income and do not vary by age cohort, that is $\tau(s,t) = \tau(t) y(s,t)$.

If as before one aggregates across all individuals, total consumption for the case of logarithmic utility and a constant-through-time mortality rate – aggregating individual consumptions from equation (8.12d) – is seen to depend on a marginal propensity to consume out of total wealth (financial wealth plus human wealth), with the propensity depending on the rate of time preference and the death rate:

$$C(t) = (\theta + \bar{p})[FW(t) + HW(t)]. \quad (8.16)$$

HW is aggregate human wealth and FW is aggregate financial wealth (holdings of domestic money, domestic bonds, equity claims on the domestic capital stock, and net foreign assets). For the more general case where utility is modeled with the CRRA function and the mortality rate is time varying, the marginal propensity to consume out of wealth is less simple than in equation (8.16). For that case, aggregate consumption still depends on the sum of financial and human wealth; the marginal propensity to consume out of wealth, however, depends not only on θ but also in a non-linear way on the IES (which need not be unity) and the entire future path of expected real interest rates and time-varying mortality rates $p(t)$.

The change in aggregate human wealth, a variable representing the present value of economy-wide labor income (adjusted for the varying ages and relative productivities of different cohorts), is given by:

$$\dot{WH}_t = \frac{d}{dt}\int_{-\infty}^{t} wh(s,t)N(s,t)ds \tag{8.17a}$$

$$= wh(t,t)b(t)N(t) + r(t)WH(t) - [Y(t) - T(t)] \tag{8.17b}$$

where T is total taxes on labor income and $Y - T$ is total disposable labor income.[26]

Equations (8.10), (8.16), and (8.17) are the key macroeconomic relationships in the modified approach taken here. Changes in the demographic composition of the population and effective labor force significantly influence the aggregate supply of labor, given the differences across age groups summarized in the age–earnings profile. On the demand side of the economy, aggregate consumption and saving behavior – and hence also wealth accumulation – are strongly influenced by the life-cycle effects of the demographic changes on human wealth.

4. OUR STRATEGY FOR TESTING THE MODIFIED APPROACH IN MULTI-COUNTRY EMPIRICAL MODELS

Multi-country macroeconomic models are inherently complex. Differences among countries or regions in any given model interact in numerous ways with the model's specifications of economic behavior, thereby making it difficult to understand the consequences of changing the model's specification for any one aspect of behavior. Since alternative competing models have significant differences in the treatment of economic behavior and of particular nations or regions, comparisons across alternative multi-country models are especially complex (Bryant et al. 1993).

To facilitate a clearer comparison of models and alternative treatments of economic behavior, the authors have constructed stylized and simplified versions of two existing macroeconomic models. Several years ago Bryant, together with Long Zhang (1996a, 1996b), developed a stylized, two-region abridgement of the IMF staff's MULTIMOD model labeled the Bryant-Multimod-2-Region model or BM2R model. More recently McKibbin has developed a stylized, two-country abridgement of the McKibbin and Wilcoxen G-Cubed models, called the McKibbin Software Group Model or MSG3 model (McKibbin and Vines 2000). These abridgements of the larger models BM2R and MSG3 are research environments within which analytical comparisons are less complex.

Our preliminary work in this project has focused on refining these two stylized models to incorporate the modified approach for treating demographic changes and their macroeconomic consequences. Each author has made changes in his stylized model along the lines sketched above. The process has required frequent interactions to identify problems and work out their solution.

A subsequent research report will provide detailed descriptions of the BM2R and MSG3 models and our incorporation into them of endogenous demographics and the bottom-up determination of labor income and human wealth resulting from the incorporation of age–earnings profiles. In this chapter, we provide only a sketch of the two abridged models and identify the most important similarities and differences among them.

The starting point for both stylized models is a set of equations describing the US economy (US for short). Then a second artificial country is created, labeled for brevity here as ZZ. The ZZ economy is an identical, mirror image of the United States. These two economies are carefully linked together with the balance-sheet and income-flow identities that would have to hold if the world were composed of only these two economies. The current-account balance and the net-foreign-asset position of the ZZ economy, for example, are the negatives of the current account and the net-foreign-asset position of the US economy. A single exchange rate exists linking the two regions' currencies and economies.

Each region consists of several types of economic agents: households, firms, a government and a central bank. The MSG3 model contains two production sectors, for energy and non-energy goods and services; output produced is a function of capital, productivity-augmented labor, energy and materials. In addition there is a sector that creates capital goods which are purchased for investment purposes by both firms and households. In the BM2R model, there exists only a single composite good produced in each country, and hence only a single type of firm and production sector; output produced is a function of capital and productivity-augmented labor.[27]

The production technologies of firms in both the MSG3 and BM2R models are represented by constant elasticity of substitution (CES) production functions. Each country's goods are imperfect substitutes; each country exports some of its production to the other country. Imports in each country are a function of income and relative prices. Agents in a given country in the models are assumed to have identical, unchanging preferences over foreign and domestic goods.

Both the MSG3 and BM2R models emphasize the forward-looking behavior of agents and presuppose that both firms and households engage in intertemporal optimization. (A partial exception in both models stems from their allowance for a fraction of consumers whose consumption is constrained by an inability to borrow and who are hence unable to smooth their consumption intertemporally.) Both models require long-run evolutions of the model economies that result in steady-state, balanced-growth equilibrium paths.

The firms in the models are characterized as price-taking entities that choose variable inputs and their levels of investment in capital so as to maximize stock-market value. Firm investments in both the MSG3 and BM2R models respond to the difference between the market value and reproduction value of the capital stock (variants of Tobin's '*q*'). The MSG3 and BM2R models, however, differ in detail in how they model this process. For example, McKibbin adapts the cost-of-adjustment models of Lucas (1967), Treadway (1969) and Uzawa (1969) and following Hayashi (1979) allows investment also to depend on current cash flow. Bryant's approach so far in the BM2R model is less explicit about adjustment costs and follows the treatment in the Mark II version of MULTIMOD (Masson et al. 1990; see also Meredith 1991).[28]

The specification of the household sector and effective labor supply in the BM2R and MSG3 models, as modified by the introduction of age–earning profiles and a bottom-up determination of labor income, is summarized above. When introducing the theoretical specifications in the stylized empirical models, we have attempted to do so in a comparable fashion. One significant difference stems from different underlying assumptions about the household's utility function, and hence about consumers' IES. The MSG3 model currently incorporates an assumption of log utility, imposing the assumption that the IES is unity. The BM2R model makes use of the CRRA utility function and would normally set the value of IES well below unity, for example at 0.5. In the simulation results for the BM2R model reported in Section 5 below the IES is assumed to be unity.[29]

Both stylized models treat labor as perfectly mobile within each of the two countries but completely immobile across the boundary separating the countries. Hence wages are equal across comparable age cohorts within each country but in general are not equal across the two countries. Over the long run, labor is inelastically supplied with respect to wages and is determined by the demographic structural equations already described. The MSG3 model allows for short-run unemployment because of sticky nominal wages, though the model converges to full employment in the long run. The BM2R model likewise forces full employment of labor and capital over the longer run.[30]

Prices are assumed to clear the goods market in the MSG3 model whereas in the BM2R model an estimated model of price stickiness is used. Thus the price stickiness in the MSG3 model is driven only by wage stickiness whereas the price dynamics in the BM2R model reflect both price and wage stickiness. This is an important difference in the models and may partially explain the larger cycle in the results for the BM2R model.

The government in each country engages in real spending on goods and services (taken as exogenous in both the BM2R and MSG3 models), can make real transfer payments to households (again taken as exogenous), raises revenues by taxing firms and households, and pays interest on its outstanding stock of debt. Both models assume that agents will not hold government debt unless per capita government debt is eventually forced to grow at a rate less than the interest rate paid on the debt. Both models for this exercise use a variant of a debt-stock targeting rule. The differences between alternative approaches for the specification of an intertemporal fiscal 'closure rule' are carefully described in Bryant and Zhang (1996a).

The monies of the two countries appear in the models because of (implicit) transaction costs. Money demands in the models depend negatively on short-term nominal interest rates and positively on the value of aggregate output. The central bank in each country is assumed to follow a policy rule that ensures long-run nominal stability of the model's behavior. The models can enforce either a money-targeting rule, a nominal-GNP-targeting rule, or a rule combining inflation targeting with real GNP targeting. These rules are explained and analysed in Bryant et al. (1993). To facilitate comparison across the models and foster simplicity, in the first stages of this joint research both the BM2R and MSG3 models assume that the central banks follow money-targeting rules.

As a starting point for simulation experiments, we use the stylized models to develop model-consistent baseline paths for the evolution of the US and ZZ economies. In the BM2R model the steady-state baseline is used whereas in the MSG3 model a baseline on the stable path to the steady state is used. Because the US and ZZ economies are identical in the baseline solutions for the two models, the exchange rate is constant over time at unity and the trade balances, current-account balances, and net-foreign-asset positions are constant at zero. Several alternative baseline solutions have been investigated. In one baseline, the birth rates and the mortality rates in both countries are held constant at 1.5 percent per year (0.015), resulting in a population that is stationary over time. In a second baseline solution, the birth rate is held constant at 2.5 percent per year (0.025) and the mortality rate remains constant at 1.5 percent per year (0.015); in that baseline the two countries' populations grow at the constant rate of 1 percent a year.[31]

The baselines typically assume that productivity growth occurs at a constant rate (2 percent per year in these initial simulations). Baseline rates of inflation are likewise assumed constant (also 2 percent per year).

We have done preliminary experiments in both models allowing for a moderate fraction (for example, one-quarter) of consumers to be borrowing-constrained. Much of the preliminary experiments (and all those reported in this chapter) have focused on the case where that fraction is zero. This issue is briefly discussed further below.

The two modeling approaches differ in a number of important theoretical and empirical respects. We have attempted to standardize as much as possible across the models without changing the core approaches of the models. For example, in this chapter the BM2R model has used a value of unity for the elasticity of substitution between labor and capital (the Cobb–Douglas case). Also for this chapter we impose a Cobb–Douglas production function rather than those that were estimated on the MSG3 model. Despite this there are still a number of important differences in the models which show up in important ways in the differences in results. What is useful from this approach is that despite having different analytical and empirical views of the world, the two models produce many broad similarities in the results.

5. PRELIMINARY SIMULATION EXPERIMENTS: A TRANSITORY DEMOGRAPHIC BULGE IN THE TWO MODELS

In the simulations reported in this section, the focus is on a transitory demographic bulge. In the first variant, the bulge occurs simultaneously and identically in both the US economy and the ZZ economy; this variant in essence is the case of a closed world economy, since both the US and ZZ model economies are identical and are subjected to an identical demographic shift. The second variant assumes that the bulge occurs in the US economy but not in the ZZ economy; the second variant highlights the transmission of effects from an economy experiencing a demographic shock to the rest of the world.

The baby bulges – strictly speaking, in this version of the model without youth dependency, the bulge of new 20-year-olds into the adult population – take the following specific form. Prior to the shock, populations are stationary; the adult birth rates and mortality rates are all equal at the rate of 0.015. In year 1 the birth rate starts gradually to rise from the baseline rate. By year 8, the birth rate has risen to 0.03, a full 1.5 percentage points higher than in the baseline. At that point, population is growing at 1.5 percent per year rather than the baseline rate of 0 percent. After persisting at the higher level of 0.03 for 5 years, the birth rate then starts gradually to fall back. By year 22, roughly two decades after the demographic shock began, the birth rate is back to the 0.015 rate and remains at that rate thereafter. Throughout the shock, the death rate remains at 0.015. Figure 8.3 gives a visual plot of these series.

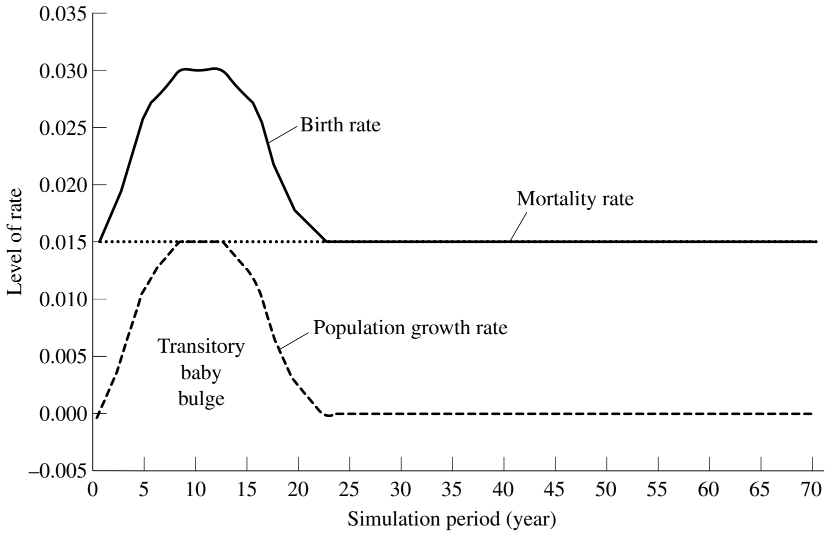


Figure 8.3 Transitory baby bulge: birth rates, mortality rates, population growth rates

For the shock variant in which the demographic bulge occurs identically in both countries, populations and the effective labor forces are eventually 22.5 percent higher than in the baseline. The labor forces adjusted for age and relative productivities rise even higher while the demographic bulge is at its peak and waning – see Figure 8.4. Figure 8.5 plots the elderly dependency ratios – see equations (8.4) and (8.5) above. Initially in the baseline, slightly more than half of the population in both countries is above the threshold age defining the elderly.[32] In the simulation, as the demographic bulge occurs, the dependency ratios fall gradually until they reach a lower level of about 42 percent of the population. After an extended period of time at that lower level, the dependency ratios start to rise back toward the steady-state baseline level. After about 70 years, the dependency ratios have risen back to their initial levels.

For the shock variant in which the demographic bulge occurs in the US only, the time paths in Figures 8.3–5 pertain only to the US. The ZZ population and effective labor force in that variant never deviate from their baseline paths. The dependency ratio in ZZ remains throughout at its initial level.

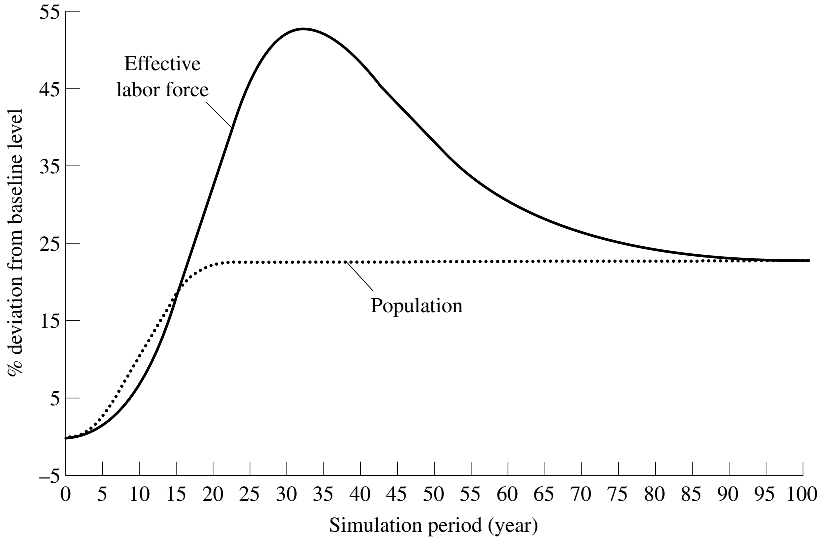


Figure 8.4 Population and labor force

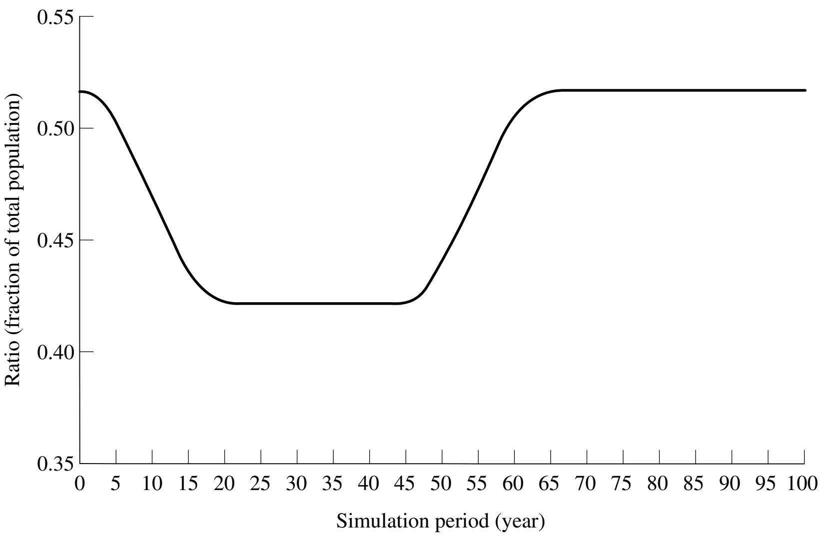


Figure 8.5 Elderly support ratios

The stylized shock studied here is analogous to the 'baby boom' experienced by several industrial nations in the sixth and seventh decades of the twentieth century. The reversal of the baby boom with the passage of time results in population ageing of the type now beginning to manifest itself in those nations.

Given both of the models' treatment of saving, investment and financial variables as forward looking, agents immediately adjust some aspects of their behavior at the onset of the shock. The numerical algorithms used to solve the models presume that agents correctly anticipate the entire future path of the demographic shock. Thus key variables like interest rates, human wealth, the market value of capital and forward-looking goods prices immediately jump to altered levels. For the shock variant in which only one of the regions experiences the demographic bulge, the nominal exchange rate and hence also the real exchange rate are significant 'jumping' variables.

The assumption that agents, at the onset of a shock, correctly anticipate the entire future paths of exogenous variables is the now-standard working assumption about expectations in models solved with rational, model-consistent expectations. Thus our use of this assumption is familiar ground. Yet the assumption is extreme. Worse, it is inherently implausible for demographic shocks such as a baby bulge that begin gradually and then also wane gradually over many years. In previous research we have shown how it is possible to modify the model-consistent-expectations assumption by phasing in 'correct expectations' with the passage of time rather than permitting expectations to be correct immediately and fully. However, for the time being we report the results here with the now familiar full model-consistent expectations.

The simulation results are shown graphically rather than in tables. All the time paths in the following figures are presented as deviations from the baseline solutions of the model (in units specified along the vertical axes of the figures). Thus if a variable has a value of zero in a figure, at that point the variable is unchanged from its baseline path.

We present the results for each model in a series of graphs. We first present a set of results for the BM2R model followed by the same results for the MSG3 model. Results are presented first for real variables, then for financial variables, and finally for international variables.

One can roughly differentiate three periods of adjustment in the results shown. Impact effects occur during the first and next few years of the shock as the population and the number of productive workers increases. (Remember that in these preliminary results we have not yet included youth dependency, so that a new member of the population is also immediately a new worker.) Given the shape of the age–earnings profile, the effective labor force increases even faster than the number of workers. Over a medium run, the (adult) birth rate stops increasing, levels out and then starts declining; during this medium run, the demographic dynamics reverse and then pick up momentum in the opposite direction. Then over a longer run, the enlarged cohorts generated by the higher birth rate pass completely through the workforce (retirement occurs gradually and continuously) and the populations eventually re-settle onto equilibrium growth paths where the growth rates are identical with those in the baseline solutions.

Figures 8.6a and 8.7a and 8.6b and 8.7b summarize the world 'closed-economy' simulation in which the demographic bulge occurs in both regions simultaneously. The figures qualified with an 'a' are those for the BM2R model; those labeled 'b' are for the MSG3 model. In the case of a world simulation, the exchange rate linking the two economies remains unchanged at unity throughout the entire simulation. The external balances of the two economies remain unchanged at zero values. The curves in Figures 8.6 and 8.7 show 'domestic' variables. These domestic outcomes are identical in both countries.

In the results for the BM2R model (Figures 8.6a and 8.7a) there are additional lines which should be ignored until the later discussion about the importance of the age–earnings profiles.

Variables such as human wealth, the market value of capital, consumption, saving and yields in asset markets are forward-looking 'jump' variables. Given the assumption of model-consistent expectations, agents correctly anticipate the future events that will unfold and hence make immediate adjustments. Households, for example, initially reduce their consumption (northwest panel of Figures 8.6a and 8.6b) and increase their savings (northeast panel) knowing that the growing population will require a larger capital stock to equip a larger labor force. In the MSG3 model, the initial jump is dampened because of the adjustment costs in capital accumulation. Despite the desire to start saving for future capital requirements, the adjustment costs in the MSG3 model lead to a postponement of this decision (a similar result is noted in McKibbin and Wilcoxen 1997). This delaying effect is much smaller in the BM2R model.

The initial movement in per capita human wealth is downward (NE panel of Figures 8.7a and 8.7b). Because of the Blanchard assumption of a positive probability of death, households partly discount the future changes in their labor incomes (relative to models in which representative agents have infinite horizons). Coincident with the other initial changes, the market value of capital relative to the existing capital stock jumps up (SE panel of Figures 8.7a and 8.7b) and real interest rates adjust downwards (NW panel of Figure 8.7a).

As the population and labor force continue to grow, per capita human

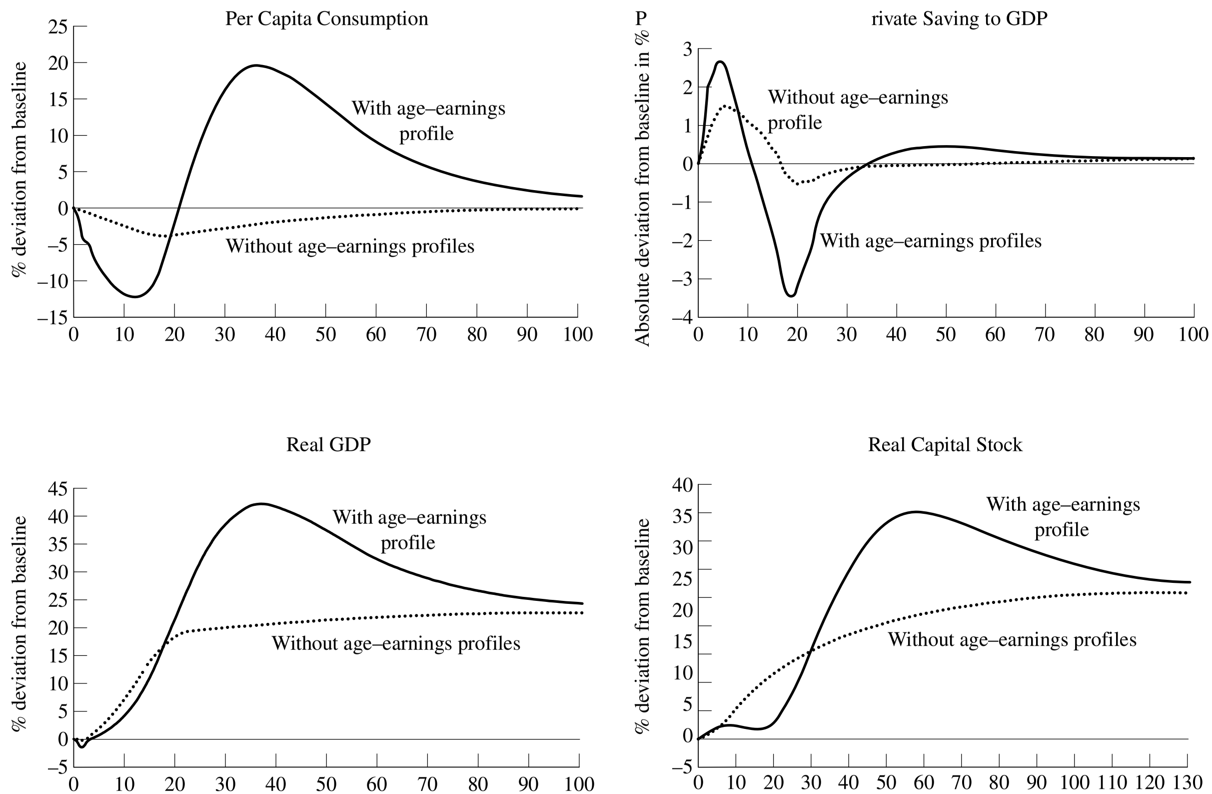


Figure 8.6a *Real implications of a baby bulge in both regions simultaneously, BM2R model*

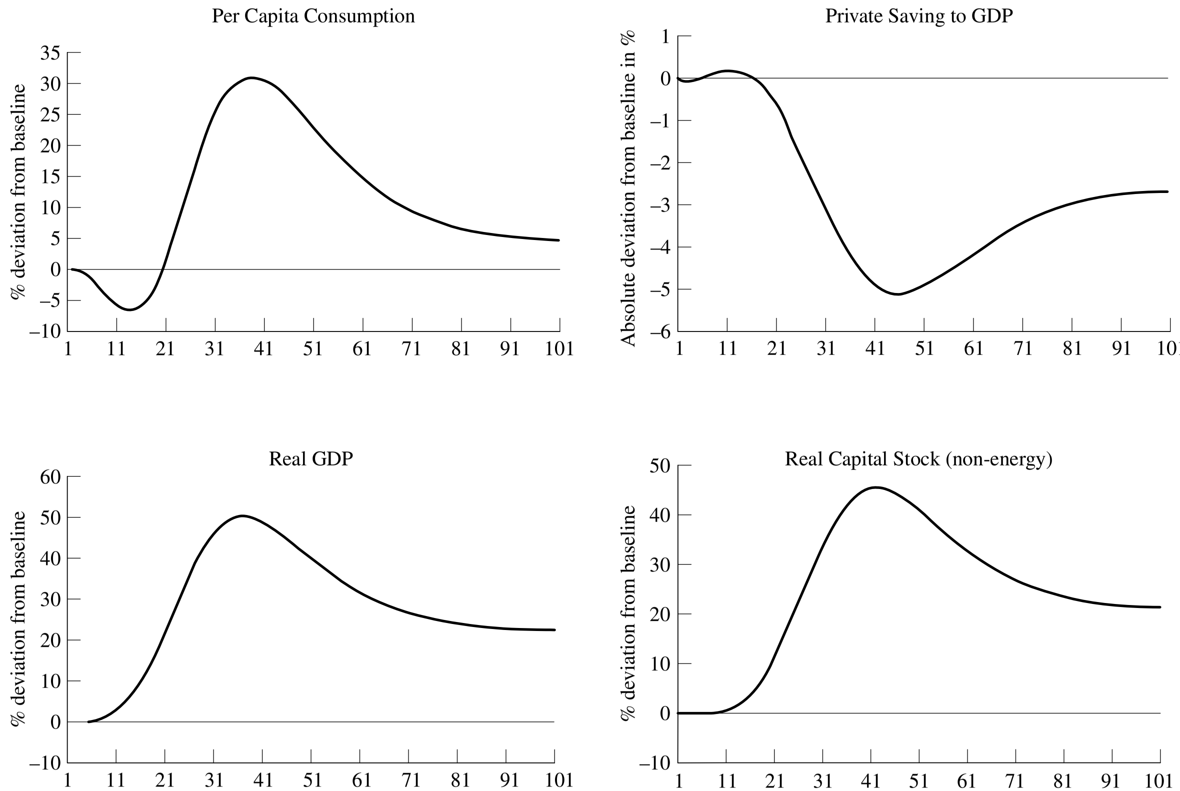


Figure 8.6b Real implications of a baby bulge in both regions simultaneously, MSG3 model

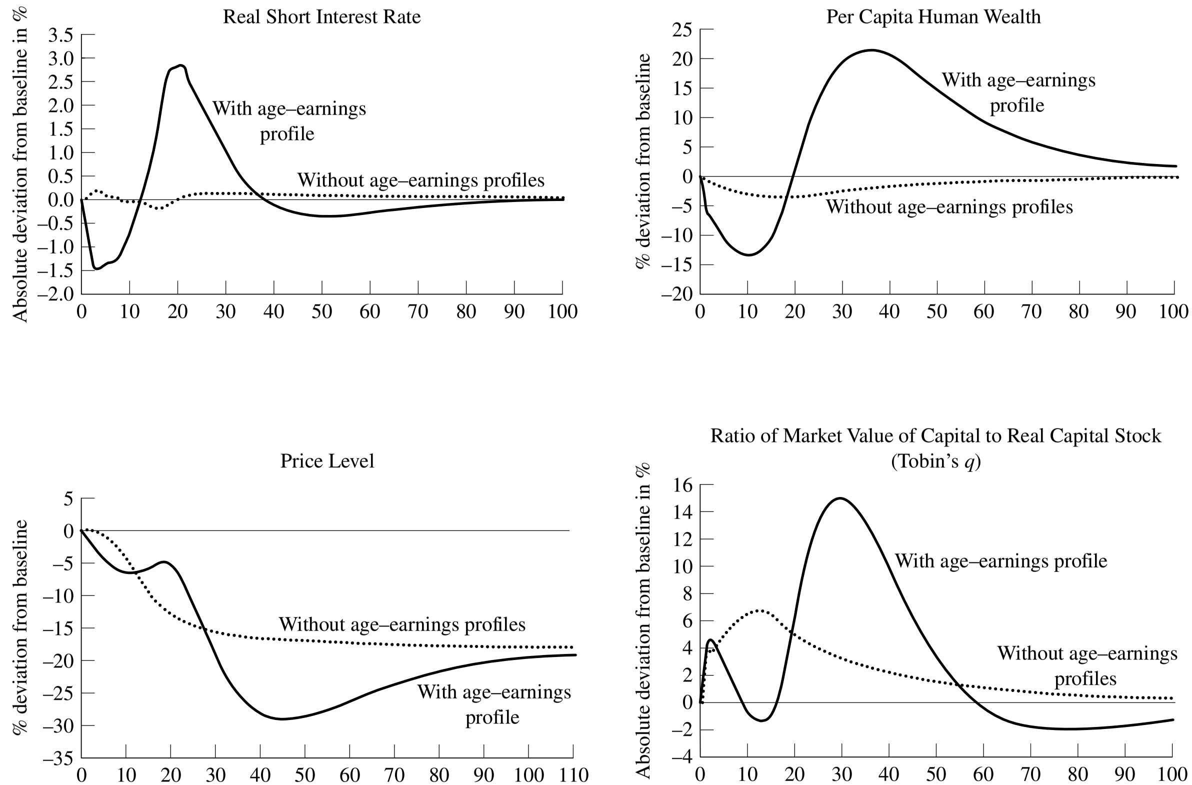


Figure 8.7a Financial implications of a baby bulge in both regions simultaneously, BM2R model

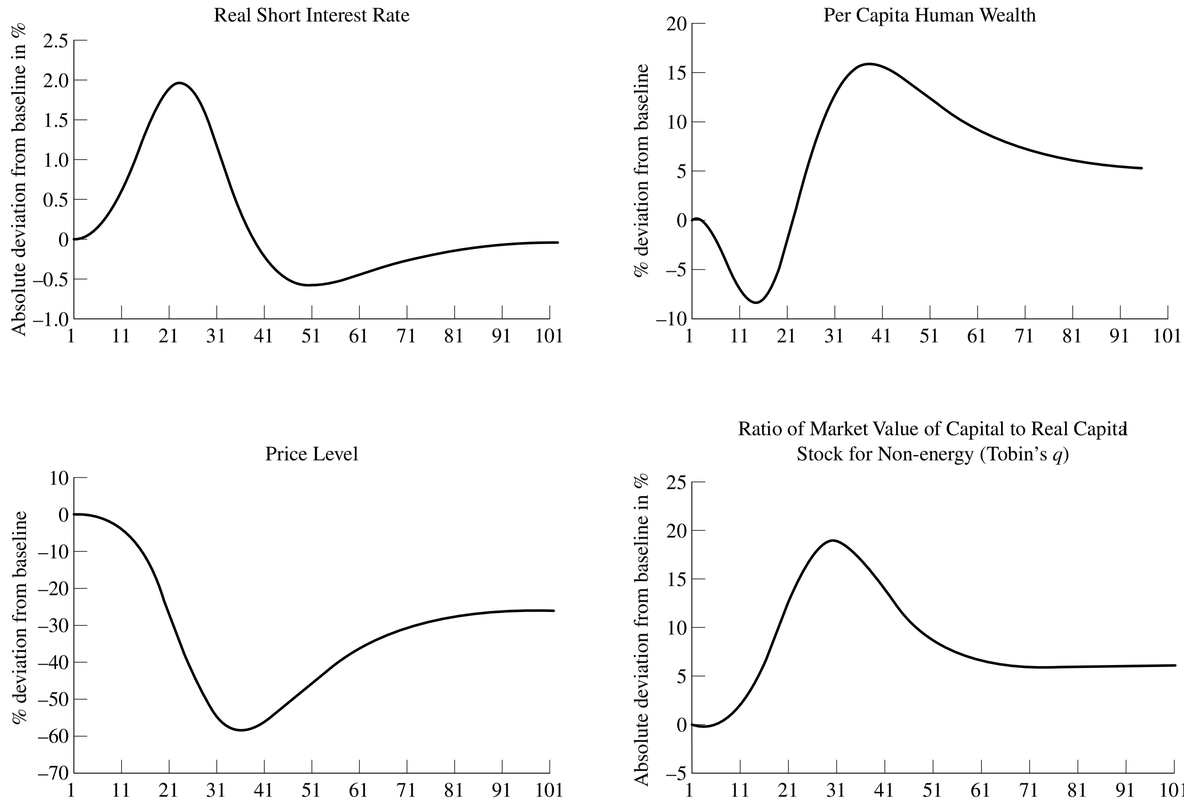


Figure 8.7b Financial implications of a baby bulge in both regions simultaneously, MSG3 model

wealth and consumption fall further. At the time when growth in the population is fastest but not increasing further, the decline in per capita consumption ceases. Then as population growth starts to fall back, fewer households with low saving rates are added to the population and households perceive that, in effect, less saving will be needed than would have occurred with continued high population growth. Per capita consumption can then begin to rise back toward and then above baseline. Because of the hump-shaped age–earning profile, the *effective* size of the labor force continues to grow rapidly long after the population itself has ceased growing (Figure 8.4). Because of the bottom-up determination of aggregate labor income, moreover, per capita human wealth begins a long and sustained rise to a level well above baseline. After its initial jump upwards, the ratio of the market value of capital to the real capital stock falls back as the population grows; that ratio, too, then reverses course and begins a sustained cyclical rise.

The cyclical movements in per capita human wealth and consumption during the years 10 to 30 and the associated cyclical behavior of other domestic macroeconomic variables are partly due to the fact that the new members of the population resulting from the (adult) baby bulge are at first relatively low savers. Speaking loosely, these younger adults are ascending the left side of the hump of their age–earning profiles (Figures 8.1 and 8.2). As the bulge cohorts reach their years of peak earnings and high savings, another inflection point is reached. Per capita human wealth and consumption in the fourth decade of the shock begin a long decline relative to baseline. Eventually, as the baby boomers become elderly, their labor income and human wealth decline and they begin to consume out of their financial wealth. Per capita consumption in the long run returns to the baseline level of the initial steady state.

The SE panel of Figures 8.6a and 8.6b show the behavior of the real capital stock. The quite modest rise of the capital stock in the initial decades are also explained by the working in the model of the age–earnings profiles and the bottom-up determination of labor income and human wealth as well as the adjustment cost approach to capital accumulation in both models. New members of society begin their lives without any financial wealth and are low savers. In the years of rapid population growth, therefore, the presence of numerous additional individuals with low saving slows down the capital accumulation that must eventually take place, leads to increases in interest rates, and explains why the capital stock declines slightly before population growth reaches its peak. Without the age–earnings profiles playing a role in the model, capital accumulation would not be so weak and interest rates would not rise so much during the years 10 to 20. Once the bulge in population has occurred and the population growth rate has returned to zero (year 22 and thereafter), the baby boomers increasingly enter into their years of high saving. Financial wealth per capita and capital accumulation itself enter a period of rapid, sustained increase. By this time interest rates are falling back, even falling below the long-run baseline level to which they will ultimately return. Eventually, as the baby boomers pass into their elderly years and the elderly dependency ratio returns back to its original baseline level, capital accumulation gradually declines somewhat and interest rates rise back to their long-run level.

The role played in the model by the presence of age–earnings profiles is central to understanding the simulation results. A useful way to see the importance of age–earnings profiles is to compare simulations in which the profiles are and are not present. Consider, therefore, the second set of curves in Figures 8.6a and 8.7a for the BM2R model (the dotted curves). These curves are generated by a model run in which everything is identical except for the fact that the three 'alpha' coefficients in the model – see equations (8.10) and (8.17) – are set equal to zero instead of their estimated values. The effect of zeroing out these coefficients is to produce a model which does *not* have the distinctive cyclical behavior that has just been discussed. It still has unrelated cohorts being born and thus is not Ricardian. For example, without the presence of the age–earnings profiles, consumption and per capita human wealth decline gradually and modestly while the demographic bulge is taking place and then gradually and slowly – without cyclical behavior – rise back to their baseline levels (NW panel of Figure 8.6a and NE panel of Figure 8.7a). In the absence of the age–earnings profiles, the capital stock would increase monotonically toward its new steady-state level and interest rates would never rise above their baseline steady-state level (SE panel of Figure 8.6a and NW panel of Figure 8.7a).

Although the incorporation of age–earnings profiles is probably the most distinctive and important feature of the model's analysis of demographic shifts, many other aspects of the models also condition the results. In an earlier version of this chapter, for example, we discussed the sensitivity of outcomes to the assumed value of the intertemporal elasticity of substitution. The simulation for the BM2R model shown here (using a value of the IES equal to unity) was compared with a simulation using an IES of 0.5. The differences between the two simulations corresponded qualitatively to what theory suggests should be true. When the IES is low rather than high, consumers are less willing to substitute future for present consumption. Consumers that are borrowing-constrained cannot be intertemporal smoothers. Per capita consumption under those conditions should therefore adjust somewhat less in response to shocks. Variables such as human wealth, interest rates and the market value of capital, on the other hand, will have to exhibit greater volatility than when the IES is higher or when no consumers are borrowing-constrained.

The preceding discussion pertains to a closed world economy with no differences among its constituent parts. Our primary analytical interest, of course, is in situations where one part of the world economy experiences different shocks and different outcomes from those occurring elsewhere. Accordingly, the remainder of the charts and the discussion focus on a situation in which the demographic bulge occurs in the US economy whereas no demographic shock occurs in the ZZ economy.

Figures 8.8a, 8.9a and 8.10a for the BM2R model and Figures 8.8b, 8.9b and 8.10b for the MSG3 model summarize the deviations of variables from the same baseline underlying Figures 8.6 and 8.7. The panels in the charts show two curves, one for US-economy variables (unbroken lines) and the other for ZZ variables (dotted). The panels showing the exchange rates (NW panel of Figures 8.10a and 8.10b) are an exception; for those panels, one of the curves is the nominal exchange rate, the other is the real exchange rate.

In the first year of the simulation following the onset of the US-only shock, the forward-looking variables in both economies exhibit immediate jumps (for the same reasons as in the worldwide closed-economy shock). In the BM2R model, in the case of interest rates and the market value of capital, the immediate adjustments are in the same direction in both economies. In the MSG3 model they move in opposite directions in both countries. For other variables such as per capita human wealth and consumption, however, the immediate adjustments are in *opposite* directions in both models. As in the closed-economy case, US households initially experience a downward jump in human wealth and consumption and raise their saving in anticipation of the fact that a growing US population will require a much larger capital stock to equip a larger labor force (Figures 8.9a and 8.9b). Households in the rest of the world, on the other hand, modestly *increase* their consumption and *reduce* their saving, knowing that per capita human and financial wealth for ZZ residents will ultimately be boosted by the demographic shock occurring in the US.

Over the shorter and early medium runs, as per capita human wealth and the market value of capital are declining in the US, a similar but much more damped cyclical movement occurs for those variables in the ZZ economy (Figures 8.9a and 8.9b). Interest rates in the rest of the world, falling by less in the initial downward adjustment than in the US, likewise begin to increase, but again by less than in the US where the population is growing. During the medium and early longer runs, ZZ variables also tend to echo the cyclical movement in counterpart US variables, but in muted degree. After the demographic bulge in the US has completely passed through the US population and effective labor force (requiring the passage of some

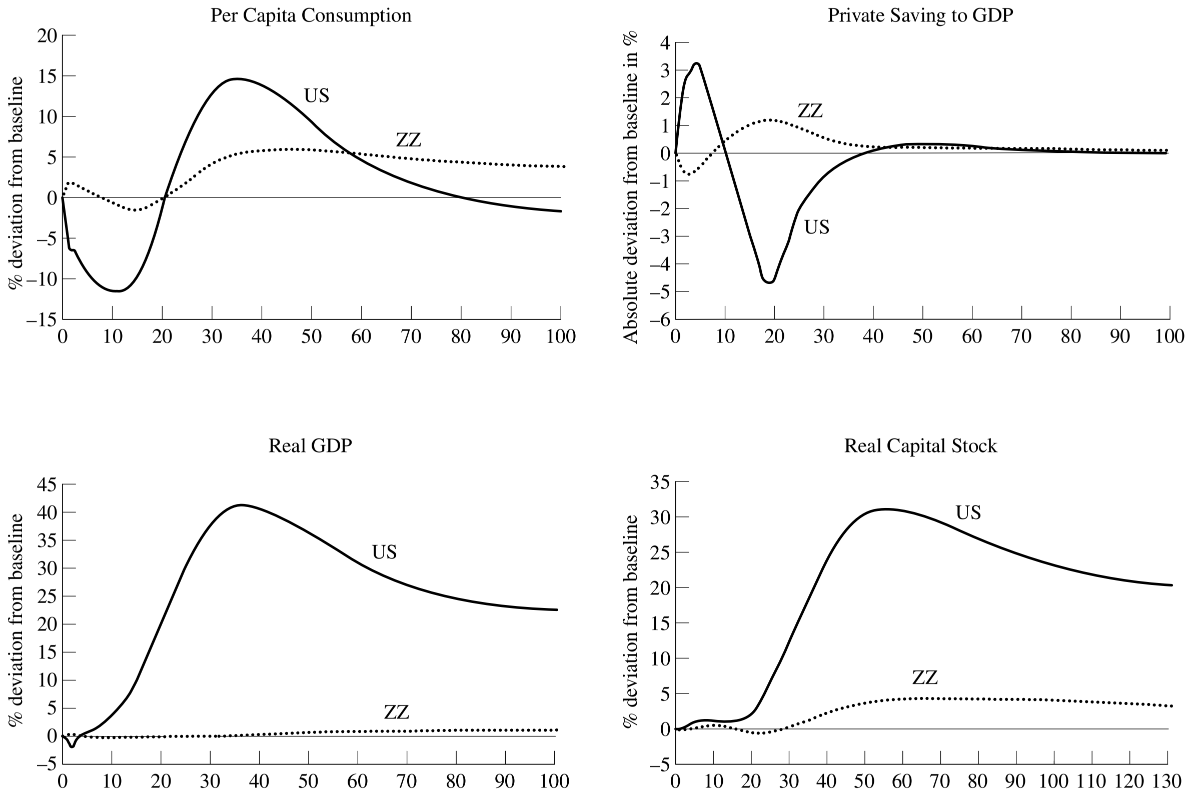


Figure 8.8a Real implications of a baby bulge in the US region only, BM2R model

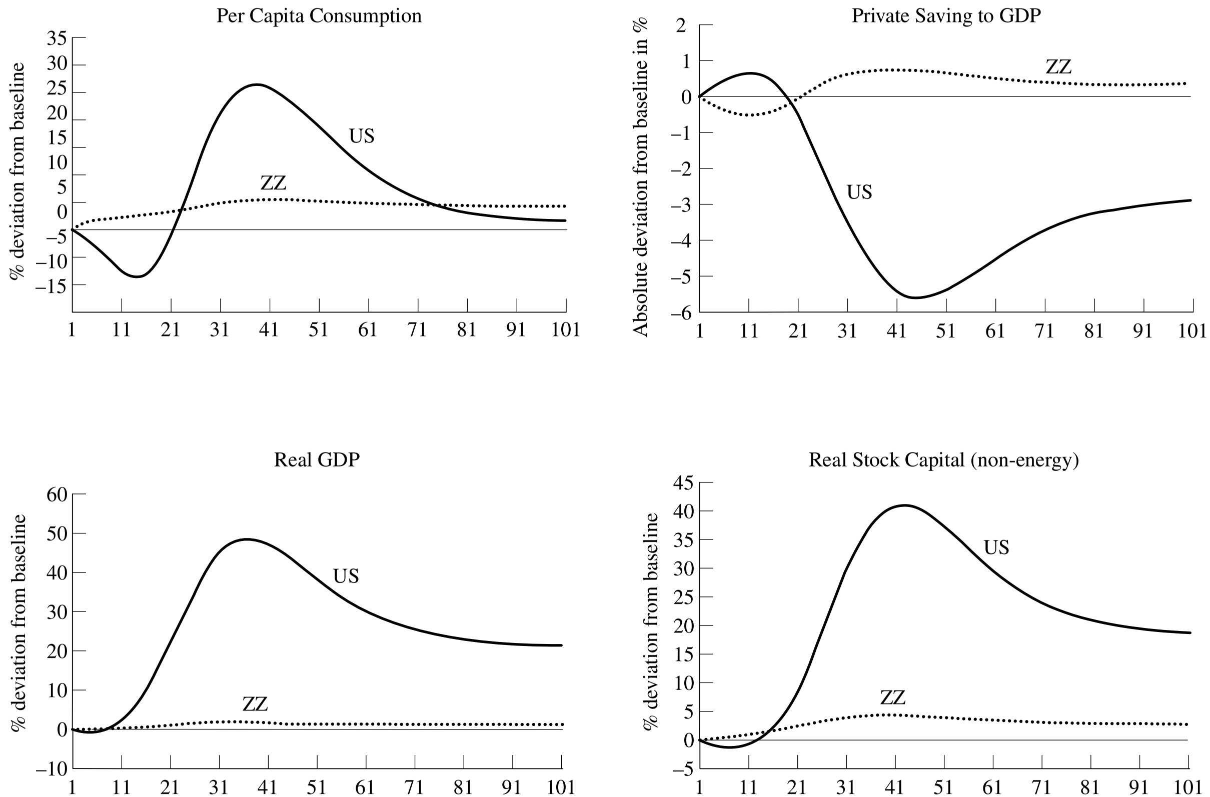


Figure 8.8b Real implications of a baby bulge in the US region only, MSG3 model

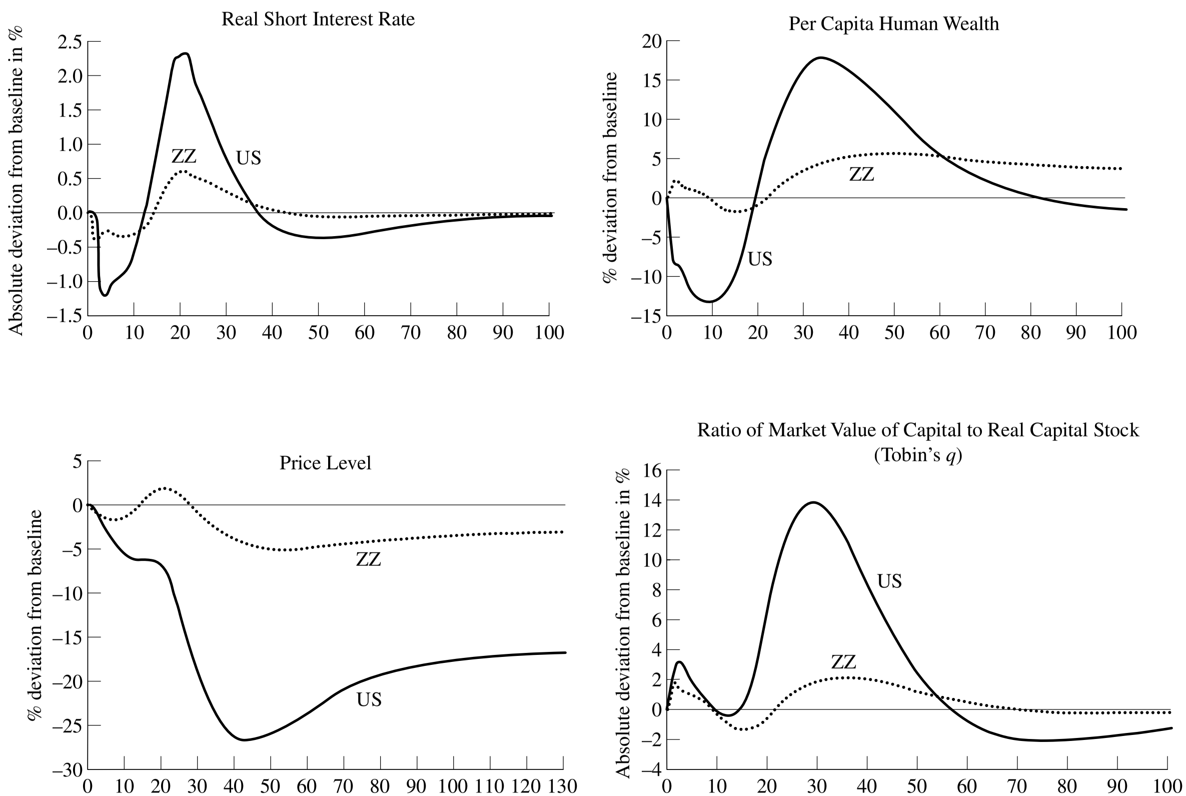


Figure 8.9a Financial implications of a baby bulge in the US region only, BM2R model

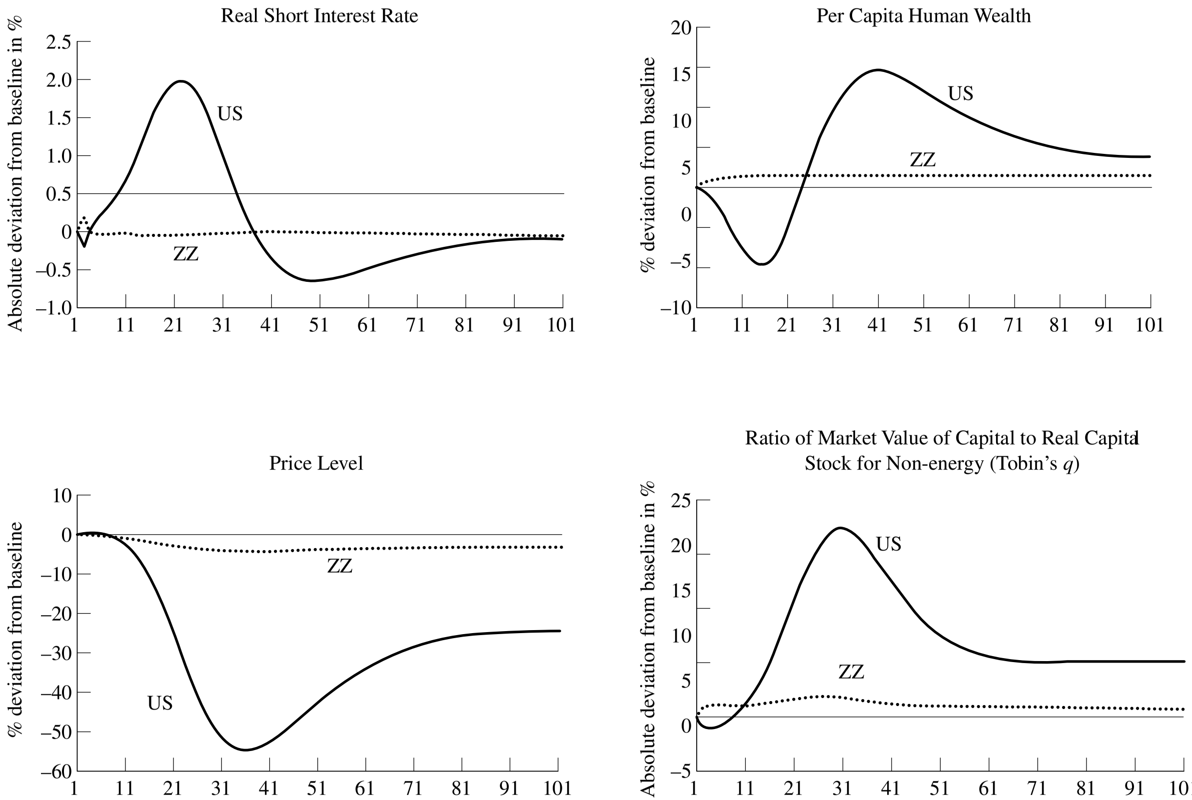


Figure 8.9b Financial implications of a baby bulge in the US region only, MSG3 model

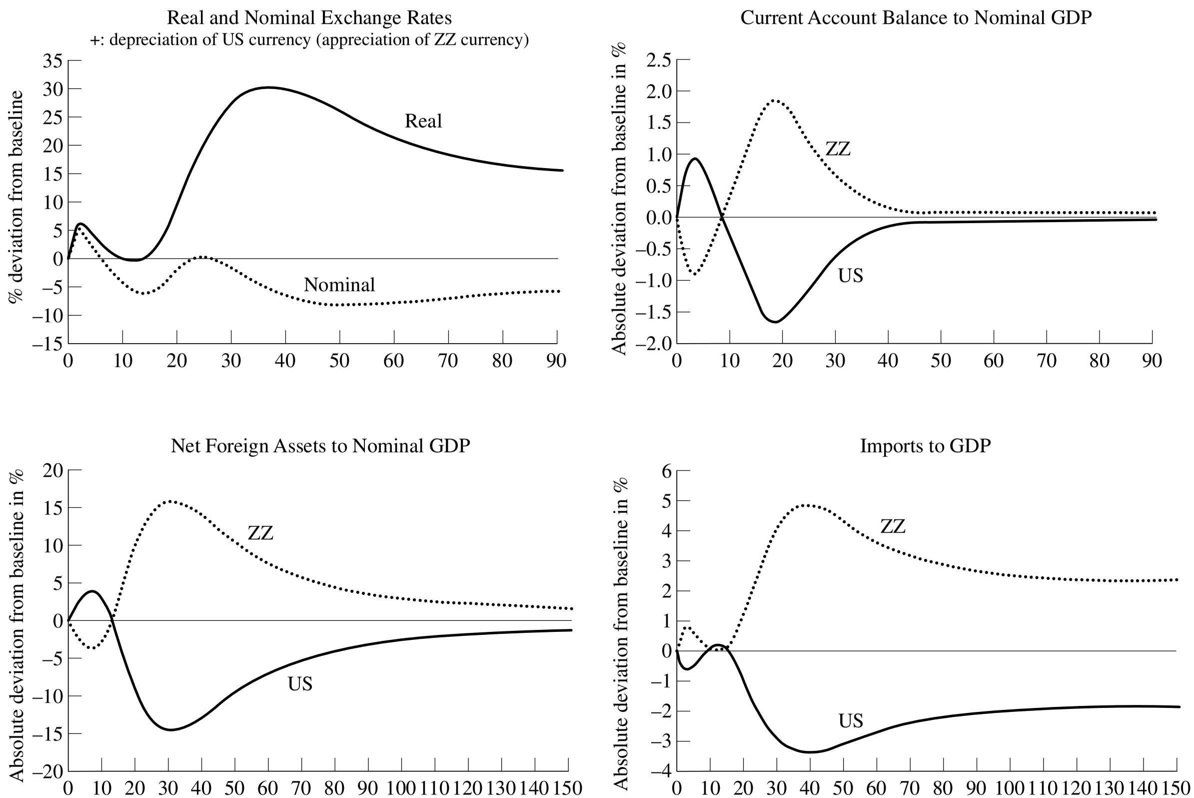


Figure 8.10a International implications of a baby bulge in the US region only, BM2R model

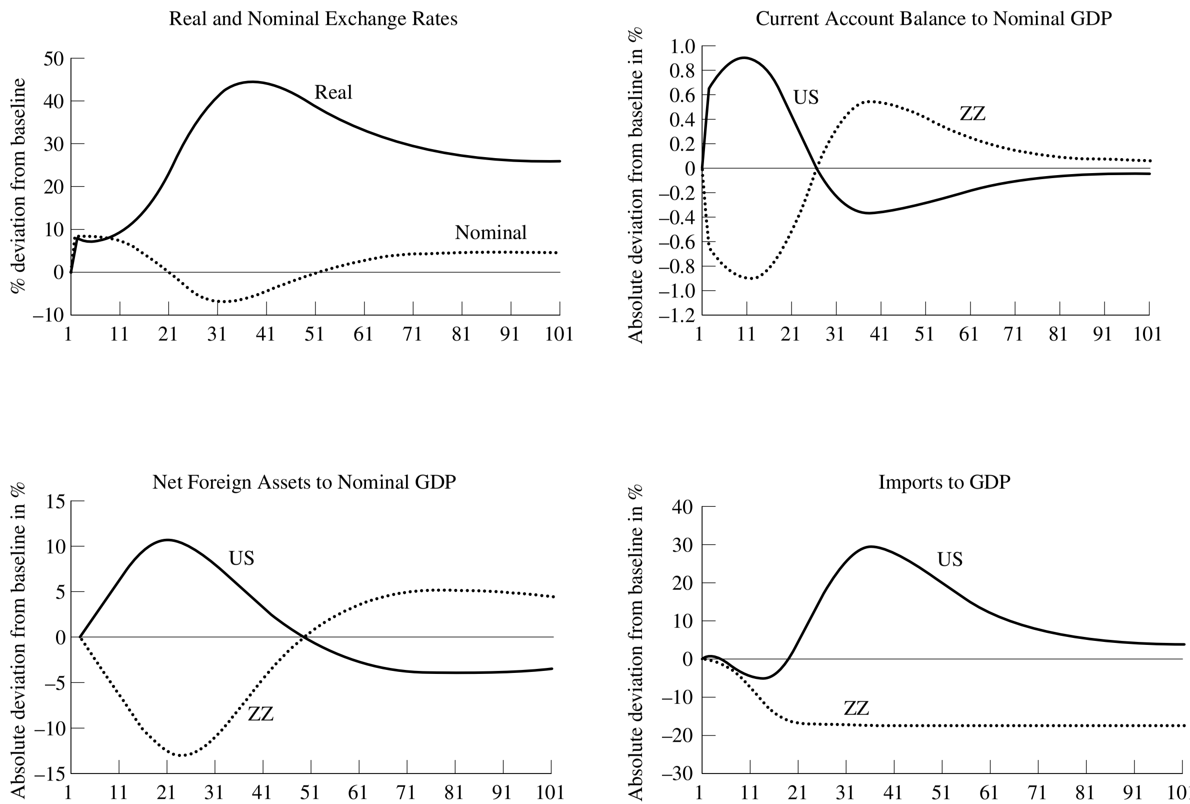


Figure 8.10b International implications of a baby bulge in the US region only, MSG3 model

75–90 years for the effective labor force – see Figure 8.4), all US and ZZ variables are beginning an approach to their eventual new steady-state levels.

One important result immediately evident in Figures 8.8a and 8.8b is that the cyclical dynamics present in the worldwide closed-economy case are characteristic of the open-economy case as well (compare with Figures 8.6a and 8.6b). This result is not surprising for the economy in which the demographic shock originates. But it is less obvious that qualitatively similar cyclical movements will occur in the rest of the world. Yet, in this model environment, the US-originating shock is strongly transmitted abroad through the exchange rate, through trade volumes and prices, and through shifts in the external balances of the countries. The exchange rate and external-sector variables reflect the cyclicality originating in the US economy. This transmission of the shock through the cyclical variation in external variables, interacting with the age–earnings profiles of foreign workers and their saving and consumption decisions, produces more muted but significant long swings in the ZZ economy as well. The macroeconomic outcomes in the foreign as well as shock-originating country are powerfully driven by the age–earnings profiles in the model (see Figures 8.10a and 8.10b).

Consider the behavior of the nominal and real exchange rates (NW panel of Figures 8.10a and 8.10b). Both the nominal and real exchange rate initially jump upward, a depreciation of the US currency (appreciation of the ZZ currency). In the BM2R model both exchange rates then move gradually back toward or even below baseline in the rest of the first decade (an appreciation of the US currency). In the MSG3 model the nominal and real exchange rates are relatively flat for the rest of the initial decade. Over the medium run, as the rate of population growth reaches its peak and then falls back toward zero, the real value of the US currency then begins a period of strong depreciation in both models. Over the longer run, the higher output of US goods resulting from the higher population and effective labor force will lower the price of US goods relative to foreign-produced goods (given the assumption that US and ZZ goods are imperfect substitutes in consumption bundles). Thus the paths of the real exchange rate and nominal exchange rate must eventually be quite different. In the long run, the US currency must exhibit a significant *real depreciation*. Given differences in the evolution of price levels in the two countries (SW panel of Figures 8.9a and 8.9b), however, the US currency in the longer run must exhibit a *nominal appreciation vis-à-vis* the ZZ currency.

The shorter-run effects of the initial depreciation of the US currency are to reduce US imports and to stimulate US exports (increase foreign imports and reduce their exports). Thus for the first few years of the simulation, US residents do not import savings from the rest of the world but rather export some of their initially higher savings abroad. The timing and magnitude of this adjustment differs across the two models but are broadly consistent. The US trade balance and current-account balance show a modest surplus in the initial years as capital is exported from the country experiencing the demographic shock (the US). As the medium-run cyclical movements in consumption, output and the capital stock occur, however, the US external balances move into deficit and foreign savings begin to flow net into the US to finance the forthcoming large buildup of the US capital stock. The US deficits bottom out (ZZ external surpluses stop increasing) around the time that population growth in the US has fallen back to its baseline growth rate of zero. From that point onward, the external balances (the ZZ external balances are of course exactly equal to those of the US but with opposite sign) commence a gradual and protracted move back toward baseline and their ultimate long-run steady-state ratios relative to nominal GDP.

Although the ratios to nominal GDP of the current-account balances of the two countries move back toward levels like those in the pre-shock steady state, the ultimate long-run ratios differ non-trivially from the baseline levels of exactly zero (NE panel of Figures 8.10a and 8.10b). The US current-account ratio moves to a permanent small negative number. The ZZ economy has a permanent surplus. The net-foreign-asset positions of the two economies (SW panels of Figures 8.10a and 8.10b) of course are the integral over time of the current-account balances.

As further evidence of how important a role the demographics and age–earnings profiles play in the behavior of the exchange rates and external balances of the countries, consider Figure 8.11. The panels of that figure repeat the curves from the simulation shown in Figures 8.8a, 8.9a and 8.10a for the BM2R model for four variables: the real exchange rate, per capita human wealth and consumption for the US, and the US ratio of net foreign assets to nominal GDP. In addition to the curves from that simulation, shown as an unbroken line, each panel contains a second curve, shown as a dotted line, obtained from a simulation identical in every respect except that the 'alpha' coefficients in the age–earnings profile are set to zero. Again it is plain that the dominant source of the cyclicality in the model's behavior is attributable to the presence of the age–earnings profiles and the bottom-up determination of labor income and human wealth.

When the age–earnings profiles are suppressed in the BM2R model, the real exchange rate after its initial jump depreciation continues gradually to depreciate slightly further. The US external balances continue in surplus until the medium run, which causes a further buildup in the US net external asset position (NE panel of Figure 8.11). Thereafter, as the population bulge ceases and the larger cohorts pass through the US population, the US

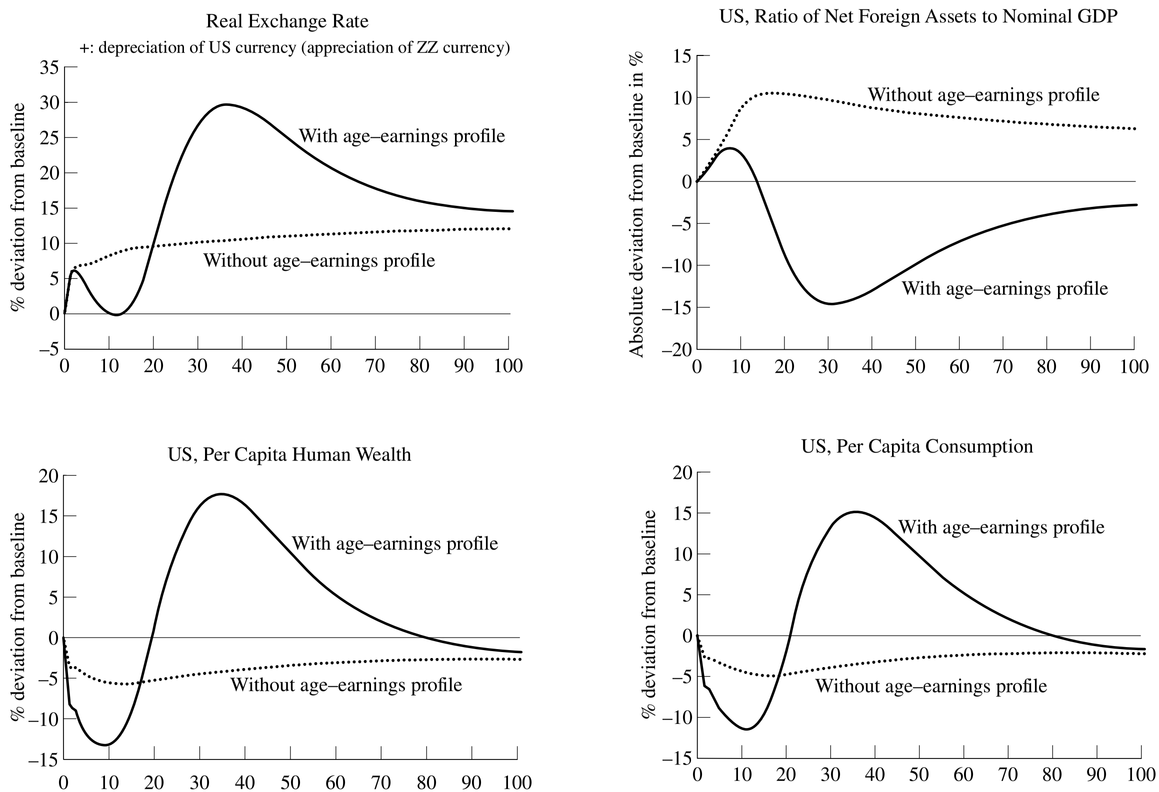


Figure 8.11 Baby bulge in US only, effects with and without age–earning profiles, BM2R model

net-foreign-asset position begins a long gradual decline. But it never turns negative, even in the very long run. In sharp contrast, when the age–earning profiles are allowed to exert their effects in the model, the initial buildup in the US net-foreign-asset position is reversed after some 6 years. The decumulation of the net-foreign-asset position is traceable to the new cohorts entering the population with their initially lower saving rates. Subsequently, when the baby boomers have ascended the left hump of their age–earnings profile, and entered their high saving years, the US current-account deficit starts to decline, which in turn eventually leads to a diminution of the negative net-foreign-asset position relative to nominal GDP. Because of the numerous years of previous current-account deficits, however, US net foreign assets remain negative even in the long-run steady state.

To give a better sense of the differences between the simulation results for the worldwide closed-economy shock and the US-only shock, Figure 8.12 shows four panels of the preceding results from the BM2R model for four US variables. The broad patterns are of course similar for the US whether the shock occurs identically in both countries or occurs only in the US. But even for the US economy, differences between the two cases are noteworthy.

The strength of the cyclical effects in the model due to the age–earnings profiles depends significantly on the demographic assumptions built into the underlying steady-state baseline. Figure 8.13 illustrates this point by comparing, for two US variables, the simulation presented in Figures 8.8a, 8.9a and 8.10a with a simulation comparable in every respect except for the fact that steady-state population growth in the alternative simulation occurs at a positive rate of 1 percent per year. Only results for the BM2R model are shown. As the charts make clear, the cyclical effects produced by the age–earnings profiles are significantly different. For example, the cyclical movements in per capita human wealth have different amplitudes. The initial jump upwards in the market value of capital is greater when the baseline population is growing steadily at 1 percent. But the upward rise relative to the actual capital stock in the medium-run years is smaller and the subsequent decline back toward the long-run steady-state level is somewhat smaller as well.

A last point stands out in the simulations when the demographic shock occurs only in the US. Welfare in the rest of the world – at least if narrowly interpreted by a measure such as real per capita total consumption (private plus government) – is *improved* in the long run by the demographic bulge in the US. (There is a transitory period in the early medium run, a bit longer than a decade, in which real per capita consumption in the rest of the world falls below baseline.) Welfare consequences in the US, again measured by the crude rubric of real per capita consumption, are complex. As the demographic bulge occurs, real per capita consumption has to fall well below

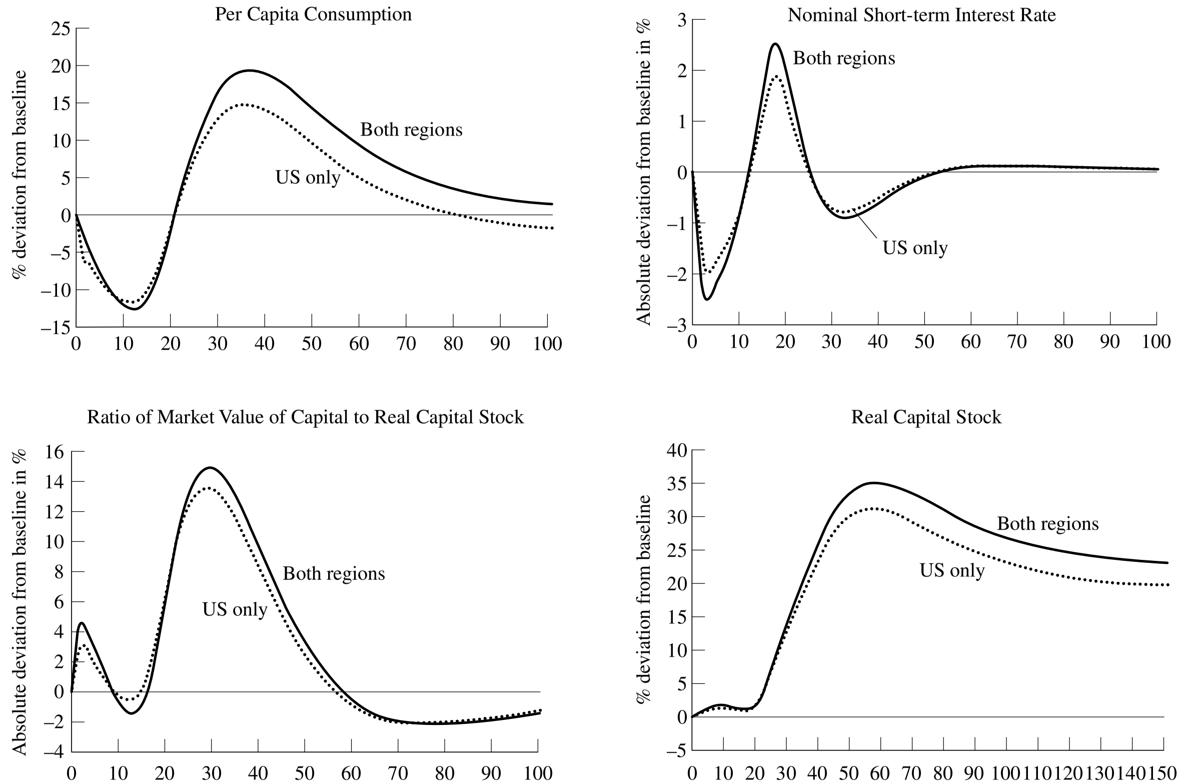


Figure 8.12 US variables, comparison of shock in both regions versus in US only

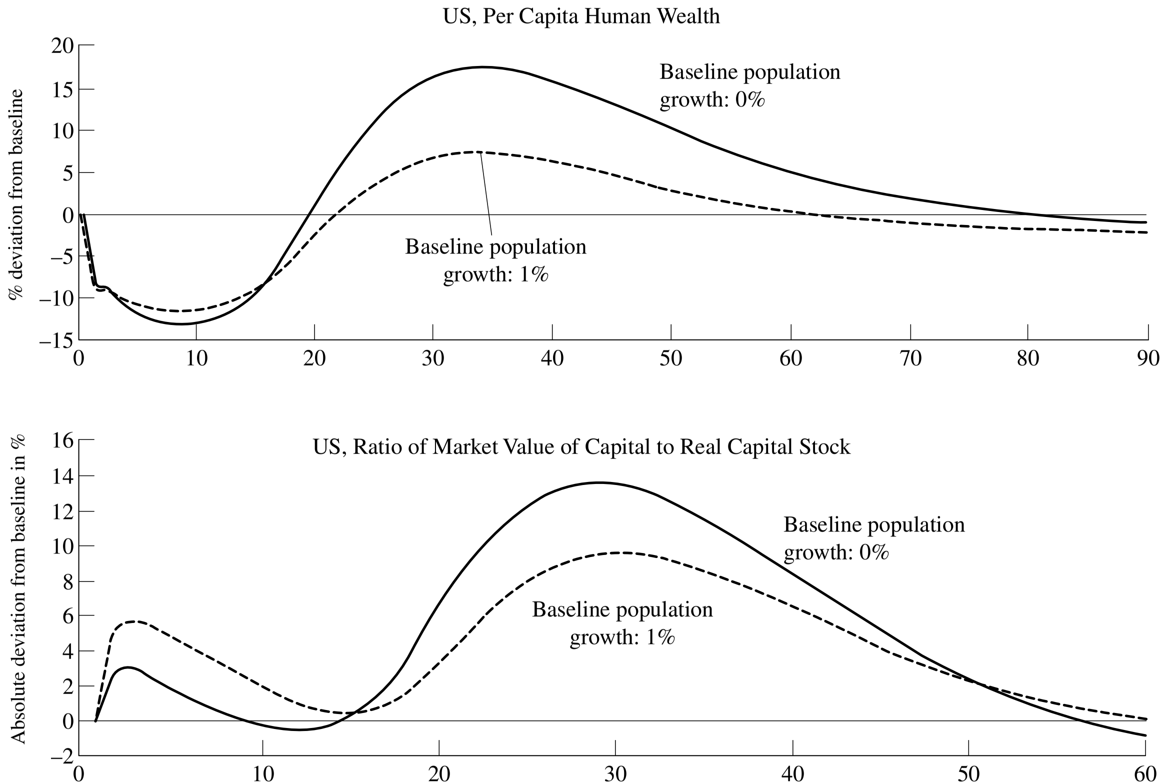


Figure 8.13 Long-run stationary population versus long-run population growth at 1% per year, baby bulge in US only

baseline. At the height of the bulge and while it is waning, on the other hand, real per capita total consumption in the US rises well above baseline. Ultimately, in the BM2R model US per capita real consumption declines to a point significantly – some 2.5 percent – below baseline. In contrast, in the MSG3 model it remains forever above baseline.

At first glance, this result may seem puzzling. Why should the average US consumer be ultimately worse off than in the steady-state baseline? The intuitive way to understand this result is that US consumers at the time the shock occurs are better off in an open economy than they would otherwise be in a closed economy, given their positive rate of time preference and given that in this Blanchard-type world they further discount the future because of an above-zero probability of death. US consumers alive at the time the demographic shock begins wish to keep up their consumption, even if their discounted gains come at the expense of future US generations. The future US generations alive many decades later have to pay investment income to foreigners because of the borrowing from foreigners that facilitated the buildup of the US capital stock required by the US population bulge. In the long run, foreigners earn net income on their permanent creditor net-foreign-asset position. Future generations of foreigners are thus better off on average in terms of their consumption per capita than they otherwise would have been. Seen in the light of these considerations, the outcomes for per capita consumptions in the two countries are not puzzling. US consumers in the shorter run have higher welfare, in the sense that the US can keep consumption per capita above the levels that would otherwise have to prevail in a closed economy. But the higher welfare in the shorter run occurs at the expense of their children and grandchildren.

6. TENTATIVE CONCLUSIONS

The preliminary findings in our joint research – the results reported here are illustrative of those findings – are promising. As in the initial efforts of Faruqee (2000a), a Blanchard-type approach with the proposed modifications for a 'bottom-up' modeling of age–earnings profiles has proven feasible and produces plausible conclusions.

Our research shows that demographic effects can be incorporated into a framework that endogenously determines all the key macroeconomic variables involved in consumption, saving and wealth accumulation. Most importantly, the framework permits study of external-sector macroeconomic variables and spillover influences on foreign countries. As illustrated in the preceding discussion, exchange rates and cross-border trade and capital flows can be powerfully influenced by underlying demographic changes.

More work is of course needed to refine and test our approach further. Children and youth dependency need to be introduced in the models. Our presumption is that the incorporation of youth dependency will have major influences on the dynamic effects of demographic change. We also plan to incorporate an explicit simplified pension-social-security system in the models. Tax revenues will be collected from working cohorts of the population and transfers will be made to elderly members. Any imbalances between revenues and transfers will result in government budget imbalances and changes in government debt stocks, which in turn will have significant macroeconomic effects. A further refinement we would like to make is a relaxation of the assumption that an economy's age–earnings profile is unchanged through time.

Despite the need for such improvements, we now feel confident that significant progress can be obtained along the lines developed in this chapter. We thus plan to continue with development of our two-region stylized models, working out identified problems and improving the models' abilities to capture key features of demographic influences on macroeconomic behavior.

Our simplified approach ignores immigration and emigration. Immigration is a quantitatively significant phenomenon for the United States and several European nations. Immigration has not been, at least not yet, a significant factor in the evolution of Japanese demographics. As of now, we do not see an easy way of remedying the drawback that our model ignores the flows of people across national borders.

In addition to the refinements already identified, we believe a great deal can be learned about the impacts of demographic change within our stylized models by undertaking sensitivity analyses. These analyses could focus on, for example, the modeling of investment decisions, the role of wage institutions in the short run, the implications of alternative assumptions about the elasticity of intertemporal substitution, the importance of alternative specifications of fiscal closure rules, and the importance of the degree of stickiness and liquidity constraints in the economy. Figures 8.7 and 8.11 above illustrate the type of insights that can be generated with such sensitivity analyses.

Eventually, our research must move to a second main stage. The two-region stylized models need to be expanded to include specific countries and to incorporate some of the main differences in macroeconomic behavior that are different across countries. The countries to receive first emphasis should be Japan and, probably, the European Union nations. Further work will also be needed on the existing equations for the United States. Ultimately, the multi-country models need to incorporate the macroeconomic behavior of several developing-country regions, again putting emphasis on the interactions between demographics and key macroeconomic variables.

The McKibbin–Wilcoxen MSG3 and G-cubed models will easily lend themselves to extension and modification along these lines. The insights obtained from the MSG3 stylized model will point the way to the required modifications. Analogously, the insights obtained from the BM2R stylized model can be readily adapted for incorporation in the IMF staff's MULTIMOD.

Though the second stage of the research should capture some of the most important differences among major countries or regions, the focus should continue to be on careful analysis of the macroeconomic interactions among the countries and regions. In particular, special attention should be given to exchange rates, current-account imbalances and net-foreign-asset positions.

The promising results of the project so far suggest a large agenda of possibilities for future work. Our approach may yield important insights about the global dimensions of demographic change well before it will be possible to implement successfully an explicit multi-cohort OLG approach in multi-country empirical models.

NOTES

* This chapter was presented by the authors at the Spring 2001 International Forum of Collaboration Projects sponsored by the Economic and Social Research Institute of the Japan Cabinet Office in Tokyo. The analysis was refined and further extended in subsequent research carried out in 2002–03 jointly with Delia Velculescu at Brookings and Hamid Faruqee at the International Monetary Fund. The authors are also grateful to Jeremy Nguyen for insightful analytical support and research collaboration. Rachel Rubinfeld, Elif Arbatli and Pablo Montagnes provided valuable research assistance.

1. Examples of researchers sympathetic to the life-cycle view of saving and consumption include Attanasio and Browning (1995) and Meredith (1995). Researchers emphasizing the importance of precautionary saving and doubting the empirical importance of the life-cycle theory include Carroll and Summers (1991), Deaton (1991, 1992), Carroll (1992, 1997), and Carroll and Samwick (1997); the papers in the international-comparison volume edited by Poterba (1994) tend to have a similar emphasis.
2. The microeconomic analyses of saving behavior have not produced a consensus explanation for these time trends. In particular, the demographic components of the simpler versions of the life-cycle hypothesis are certainly not providing the explanation. On the relevance of life-cycle models, see among others Bosworth et al. (1991), Hubbard et al. (1994), Deaton and Paxson (1997) and Lusardi (1998).
3. Here is another example: if the provisions in the government pension program specify a fixed age for retirement, increases in life expectancy may then increase the incentives of a worker to save through channels outside the pension system.
4. Recent references on liquidity-constrained consumption behavior include Zeldes (1989), Deaton (1991), Attanasio and Browning (1995) and Berloffa (1997). The buffer-stock models of Carroll (1992, 1994, 1997) emphasize the precautionary-saving motive; in those models, there is an implicit constraint on intertemporal smoothing by consumers,

but it does not arise because of an explicit constraint on consumers' ability to borrow against future income.

5. References on the role of bequests in saving include Davies (1981), Menchik and David (1983), Abel (1985), Bernheim et al. (1985), Bernheim (1991), Borsch-Supan and Stahl (1991), Altonji et al. (1992), Laitner and Juster (1996) and Wilhelm (1996).
6. Deaton suggests that a preferable taxonomy of intertemporal choice models would lump together the (formal) permanent-income and (conventional) life-cycle models and differentiate those models sharply from approaches incorporating liquidity constraints and/or precautionary saving. For more recent surveys of the consumption–saving–wealth-accumulation literature, see Muellbauer and Lattimore (1995), Browning and Lusardi (1996), Deaton and Paxson (1997) and Attanasio (1999). Muellbauer and Lattimore (1995) discuss differences between 'Euler-equation approaches' versus 'solved out consumption functions', arguing that the former have been overemphasized relative to the latter in recent years.
7. The effects on the private saving rate of increasing life expectancy could work in the opposite direction, leading households to save *more* during their working years to be able to sustain their consumption over a longer retirement period. (For this effect to be important, of course, the life-cycle motive has to be an important determinant of saving behavior.)
8. Models with life-cycle behavior seem an especially natural analytical environment in which demographics could influence consumption and saving. But as has been pointed out to us by Delia Velculescu, models other than those with explicit life-cycle elements may also have links between demographic changes and private saving behavior. For example, Cutler et al. (1990) and Velculescu (2000) use a Ramsey growth model which allows for dependency and their model predicts significant changes in saving when the proportion of young and old dependents in the population changes.
9. Cross-country empirical studies found a positive relationship between faster growth and higher saving; see, for example, Modigliani (1970). Subsequent empirical work with microeconomic data questions whether the correlation observed in the cross-country comparisons can be correctly attributed to life-cycle saving behavior; see, for example, Paxson (1996) and Deaton and Paxson (1997).
10. A strand in the literature on development economics was sensitive to these issues; see, for example, Coale and Hoover (1958), Mason (1987, 1988), and Taylor and Williamson (1994). The builders of empirical macroeconomic models, however, did not try to incorporate this work. Moreover, the development-economics literature did not directly address the details of how to adapt the consumption–wealth specifications in general-equilibrium macro models. The development-economics literature on population dynamics and saving is reviewed in recent papers by Bloom and Williamson (1997) and Higgins and Williamson (1997). Paxson (1996) and Deaton and Paxson (1997, 1998a, 1998b) are recent studies. Another important exception is the paper by Cutler et al. (1990).
11. In the microeconomic empirical research on consumption, some conclusions have been found to be sensitive to whether the consuming agents are defined as 'households' or as 'individuals'. For discussion, see Deaton and Paxson (1998a, 1998b), who use data for Taiwan and construct life-cycle saving profiles defined for cohorts of individuals rather than cohorts of households; their analysis shows stronger demographic effects on age profiles of saving for the individual than the household definition of cohorts.
12. Illustrations of OLG models which directly tackle the issue of different age cohorts and their saving decisions include Auerbach and Kotlikoff (1987), Ríos-Rull (1996, 2001), Storesletten (1995, 2000), Fougère and Mérette (1997a, 1997b, 1998b) and Brooks (1998).
13. See also Masson (1992) and Masson et al. (1995, 1998).
14. In 1999–2000, Fair and Robin Brooks were experimenting with further extensions of that research, including applications of it to the equations in the Fair multi-country model.
15. Meredith circulated this update of his table, which adds several papers not included in

Meredith (1995), at the July 1998 workshops. A similar update table is given in the OECD study by Turner et al. (1998, Table 2, p. 46). In their update table, the OECD staff contrast the macroeconomic studies with the microeconomic evidence, the latter tending to find near-zero or very low effects on the saving rate from changes in dependency ratios.

16. The simplifying assumption is of course at variance with real life. Blanchard himself pointed out that the evidence on mortality rates suggests low and approximately constant probabilities of death from, say, ages 20 to 40; thereafter mortality rates in real life do rise with age (sometimes modeled by 'Gomperty's Law' as in Wetterstrand 1981), reaching rates (in the United States) in the neighborhood of 16 percent by age 80 and 67 percent by age 100. Model agents might be interpreted as families rather than individuals, and p would then be interpreted as the probability that the family ends (because, for example, members of the family die without children or current members of the family have no bequest motive). The assumption of a constant-throughout-life p is less inconsistent with real life if agents are interpreted as families. In subsequent papers reporting on our research, we shall describe a model with youth dependency in which children have a different (age-invariant) probability of death from the (age-invariant) probability for adults. Hamid Faruqee has gone still further and shown how it is possible to modify the simplifying assumption that all adults are subject to the same age-invariant probability of death.
17. The Blanchard–Yaari insurance-market assumption also requires a supplementary assumption that individuals are not motivated to leave bequests to survivors and cannot go into debt so as to die with negative bequests.
18. The marginal propensity to consume is more complex if some of the assumptions in the original basic Blanchard model are relaxed (as Blanchard himself showed). In particular, if the intertemporal elasticity of substitution (IES) differs from unity and if the probability of death is time-varying (but still age invariant), the marginal propensity to consume depends on the IES and on the entire sequence of future expected values of interest rates and future mortality rates. For the case when the IES is assumed to be unity (the case of logarithmic utility) and when the age-invariant mortality rate is assumed to be constant through time, the marginal propensity to consume out of wealth reduces to the simple form of a constant, the sum of the time preference rate and the mortality rate.
19. For analytical simplicity, youth dependents were ignored. Individuals were assumed to be born at age 20 and assumed to immediately start work in the labor force. See below for further comments about youth dependency.
20. We follow Faruqee (2000b) closely in the remainder of the exposition that follows.
21. Faruqee (2000a, 2000b) observes that the assumption that old members of the adult population die at the same frequency as young adult members has the undesired effect of overstating the share of the elderly in the adult population. We roughly compensate for this problem by postulating an adult 'birth' rate and a mortality rate that are somewhat above the steady-state levels that would otherwise seem appropriate.
22. The threshold age in actual life years would be 65 if one thinks of model agents as being born at the age of 0 rather than at age 20. As discussed above, there is no discontinuity at 'retirement' *per se*. Agents older than the threshold elderly age thus still continue to receive some labor income. But as agents get older, after passing the years of their peak earnings their income declines continuously toward zero.
23. The Japanese data in Figure 8.1 and US data in Figure 8.2 suggest that this assumption is warranted (see, however, the further discussion at the end of the chapter about relaxing this assumption in future work).
24. To ensure that the sum of the exponential terms portrays a reasonable profile and that the effective amount of labor supplied is always initially increasing (when $t = s$), the following restriction on the five parameters must also hold: $a_1 > [\alpha_3 - \alpha_2(\alpha_3 - \alpha_2)]/(\alpha_3 - \alpha_1)$.
25. In effect, agents in the model contract with the insurance company to transfer all their wealth (positive or negative) to the insurance company at their death. Given wealth in the current period, w_t, an agent receives an amount from the insurance company equal to $p_t w_t$, if still alive but pays the insurance company w_t if he/she dies.
26. Equation (8.17b) shows that the incremental change in the stock of aggregate human

wealth at time t is influenced by the additional human wealth of the newest generation born at time t, that is by $wh(t,t)$. The shape of the labor–earnings profile – embodied in the five parameters a_1, a_2, $\propto_1$, $\propto_2$ and $\propto_3$ – has a critical influence on the behavior through time of $wh(t,t)$ and hence of aggregate human wealth, $HW(t)$.

27. The full MULTIMOD model of the IMF staff does distinguish oil from other goods. Some versions have specified non-oil commodities as a separate good. To simplify, Bryant removes these distinctions from his abridgement.
28. Adjustment costs for investment in capital are modeled explicitly in the Mark III version of MULTIMOD (Laxton et al. 1998) and Bryant plans to amend the BM2R model accordingly in future work.
29. In an earlier draft of this chapter (the June 2001 version, presented in Helsinki, Finland), some sensitivity analysis was undertaken on this parameter in the BM2R model, showing that alternative assumptions can make a significant quantitative difference. Further sensitivity tests on the importance of the IES will be included in subsequent papers in this project. In the early research, the MSG3 and BM2R baseline solutions have differed somewhat in their assumptions about the rate of consumers' time preference, corresponding to their sometimes different assumptions about the elasticity of intertemporal substitution.
30. The current version of the BM2R model follows the Mark II version of MULTIMOD. Capacity utilization can differ in the short run from long-run full use of capacity. But the model does not explicitly include wages and employment, and hence does not explicitly track unemployment. The authors conjecture that some of the differences in simulation results between the two stylized models are attributable to the different specifications of the wage and price sectors in the models; we hope in future work to focus on these differences.
31. This second baseline underpins one of the set of curves shown in Figure 8.11.
32. In an economy experiencing steady-state balanced growth with a positive rather than a zero rate of growth in the population, the steady-state value of the elderly dependency ratio is well below this ratio. For example, a baseline generated by the model with a sustained 1 percent per year growth in population generates a steady-state elderly dependency ratio of approximately 0.33.

REFERENCES

Abel, Andrew B. (1985), 'Precautionary saving and accidental bequests', *American Economic Review*, **75**, 777–91.

Altonji, Joseph, Fumio Hayashi and Laurence Kotlikoff (1992), 'Is the extended family altruistically linked? Direct tests using micro data', *American Economic Review*, **82**, December, 1177–98.

Attanasio, Orazio P. (1999), 'Consumption demand', in John B. Taylor and Michael Woodford (eds), *Handbook of Macroeconomics*, vol. 1B, Amsterdam: North Holland.

Attanasio, Orazio P. and Martin Browning (1995), 'Consumption over the life cycle and over the business cycle', *American Economic Review*, **85**, December, 1118–37.

Auerbach, Alan J. and Laurence J. Kotlikoff (1987), *Dynamic Fiscal Policy*, Cambridge, MA: Cambridge University Press.

Berloffa, Gabriella (1997), 'Temporary and permanent changes in consumption growth', *Economic Journal*, **107**, March, 345–58.

Bernheim, B. Douglas (1991), 'How strong are bequests motives? Evidence based on estimates of the demand for life insurance and annuities', *Journal of Political Economy*, **99**, October, 899–927.

Bernheim, B. Douglas, Andrei Schleifer and Lawrence Summers (1985), 'The strategic bequest motive', *Journal of Political Economy*, **93**, 1045–75.
Blanchard, Olivier J. (1985), 'Debt, deficits, and finite horizons', *Journal of Political Economy*, **93**, 223–47.
Bloom, David E. and Jeffrey G. Williamson (1997), 'Demographic change and human resource development', in Jeffrey Sachs and David Bloom (eds), *Emerging Asia: Changes and Challenges*, Manila: Asian Development Bank, pp. 141–97.
Borsch-Supan, Axel and Konrad Stahl (1991), 'Life-cycle savings and consumption constraints', *Journal of Population Economics*, **4**, 233–55.
Bosworth, Barry P., Gary Burtless and John Sabelhaus (1991), 'The decline in saving: some microeconomic evidence', *Brookings Papers on Economic Activity*, 183–256.
Brooks, Robin Jermyn (1998), 'Asset market and savings effects of demographic transitions', PhD dissertation, Yale University, September 1998.
Browning, Martin and Annamaria Lusardi (1996), 'Household saving: micro theories and micro facts', *Journal of Economic Literature*, **34** (4), December, 1797–855.
Bryant, Ralph C., Peter Hooper and Catherine L. Mann (1993), *Evaluating Policy Regimes: New Research in Empirical Macroeconomics*, Washington, DC: Brookings Institution.
Bryant, Ralph C. and Warwick J. McKibbin (1998), 'Issues in modeling the global dimensions of demographic change', Brookings Discussion Paper in International Economics No. 141, Washington, DC: Brookings Institution, December.
Bryant, Ralph C. and Long Zhang (1996a), 'Intertemporal fiscal policy in macroeconomic models: introduction and major altematives', Brookings Discussion Paper in International Economics No. 123, Washington, DC: Brookings Institution, June.
Bryant, Ralph C. and Long Zhang (1996b), 'Alternative specifications of intertemporal fiscal policy in a small theoretical model', Brookings Discussion Paper in International Economics No. 124, Washington, DC: Brookings Institution, June.
Buiter, Willem H. (1988), 'Death, birth, productivity growth and debt neutrality', *Economic Journal*, **98**, June, 279–93.
Carroll, Christopher D. (1992), 'The buffer stock theory of saving: some macroeconomic evidence', *Brookings Papers on Economic Activity*, **2**, 61–135.
Carroll, Christopher D. (1994), 'How does future income affect current consumption?', *Quarterly Journal of Economics*, **109**, 111–47.
Carroll, Christopher D. (1997), 'Buffer stock saving and the life cycle/permanent income hypothesis', *Quarterly Journal of Economics*, **112**, 1–57.
Carroll, Christopher D., Jody R. Overland and David N. Weil (1997), 'Comparison utility in an endogenous growth model', *Journal of Economic Growth*, **2**, 339–67.
Carroll, Christopher D., Jody R. Overland and David N. Weil (2000), 'Saving and growth with habit formation', *American Economic Review*, **90,** June, 341–55.
Carroll, Christopher D. and Andrew A. Samwick (1997), 'The nature of precautionary wealth', *Journal of Monetary Economics*, **40**, September, 41–71.
Carroll, Christopher D. and Lawrence Summers (1991), 'Consumption growth parallels income growth: some new evidence', in B. Douglas Bernheim and John B. Shoven (eds), *National Saving and Economic Performance*, Chicago and London: University of Chicago Press for the National Bureau of Economic Research, pp. 305–43.
Carroll, Christopher D. and David N. Weil (1994), 'Saving and growth: a reinter-

pretation', *Carnegie-Rochester Conference Series on Public Policy*, **40**, June, 133–92.
Coale, Ansley J. and Edgar M. Hoover (1958), *Population Growth and Economic Development in Low-income Countries*, Princeton, NJ: Princeton University Press.
Cutler, David M., James M. Poterba, Louise M. Sheiner and Lawrence H. Summers (1990), 'An aging society: opportunity or challenge?', *Brookings Papers on Economic Activity*, **1**, 1–73.
Davies, James (1981), 'Uncertain lifetimes, consumption and dissaving in retirement', *Journal of Political Economy*, **89**, 561–78.
Deaton, Angus S. (1991), 'Saving and liquidity constraints', *Econometrica*, **59**, 1121–42.
Deaton, Angus S. (1992), *Understanding Consumption*, Oxford: Oxford University Press.
Deaton, Angus S. and Christina H. Paxson (1997), 'The effects of economic and population growth on national saving and inequality', *Demography*, **34**, February, 97–114.
Deaton, Angus S., and Christina H. Paxson (1998a), 'Saving and growth: another look at the cohort evidence', Working Paper, 182, Princeton, NJ: Princeton University, Woodrow Wilson School – Development Studies.
Deaton, Angus S. and Christina H. Paxson (1998b), 'Growth, demographic structure, and national saving in Taiwan', Working Paper, 183, Princeton, NJ: Princeton University, Woodrow Wilson School – Development Studies.
Fair, Ray C. and Kathryn M. Dominguez (1991), 'Effects of the changing U.S. age distribution on macroeconomic equations', *American Economic Review*, **81**, December, 1276–94.
Faruqee, Hamid (2000a), 'Population aging and its macroeconomic implications', Chapter I in *Japan: Selected Issues*, IMF Staff Country Report No. 00/144, Washington, DC: International Monetary Fund, November.
Faruqee, Hamid ([2000b], 2002), 'Population aging and its macroeconomic implications: a framework for analysis', IMF Working Paper WP/02/16, Washington, DC: International Monetary Fund, January, Original draft October 2000.
Faruqee, Hamid, Douglas Laxton and Steven Symansky (1997), 'Government debt, life-cycle income and liquidity constraints: beyond approximate Ricardian equivalence', *IMF Staff Papers*, **44**, September, 374–82.
Feldstein, Martin S. (1980), 'International differences in social security and saving', *Journal of Public Economics*, **14**, October, 225–44.
Fougère, Maxime and Marcel Mérette (1997a), 'Population ageing and the current account in selected OECD countries', Department of Finance, Government of Canada, Working Paper, December.
Fougère, Maxime and Marcel Mérette (1997b), 'Population ageing and economic growth in seven OECD countries', Department of Finance, Government of Canada, Working Paper, December.
Fougère, Maxime and Marcel Mérette (1998a), 'Economic dynamics of population ageing in Canada: an analysis with a computable overlapping-generations model', Department of Finance, Government of Canada, Working Paper, May.
Fougère, Maxime and Marcel Mérette (1998b), 'Population ageing, intergenerational equity and growth: analysis with an endogenous growth, overlapping-generations model', Department of Finance, Government of Canada, Working Paper, June.

Friedman, Milton (1957), *A Theory of the Consumption Function*, Princeton, NJ: Princeton University Press.

Graham, J.W. (1987), 'International differences in saving rates and the life cycle hypothesis', *European Economic Review*, **31**, 1509–29.

Gruber, Jonathan and David A. Wise (eds) (1998), *Social Security Programs and Retirement Around the World*, Chicago: University of Chicago Press.

Hayashi, Fumio (1979), 'Tobin's marginal q and average q: a neoclassical interpretation', *Econometrica*, **50**, 231–24.

Higgins, Matthew and Jeffrey G. Williamson (1997), 'Age structure dynamics in Asia and dependence on foreign capital', *Population and Development Review*, **23**, June, 261–93.

Horioka, Charles Yuji (1986), 'Why is Japan's private saving rate so high?', Unpublished MS, Washington, DC: International Monetary Fund, June.

Horioka, Charles Yuji (1991), 'The determination of Japan's saving rate: the impact of the age structure of the population and other factors', *Economic Studies Quarterly*, **42**, September, 237–53.

Hubbard, R. Glenn, Jonathan Skinner and Stephen P. Zeldes (1994), 'The importance of precautionary motives in explaining individual and aggregate saving', *Carnegie Rochester Conference Series on Public Policy*, **40**, June, 59–125.

Koskela, E. and M. Viren (1989), 'International differences in saving rates and the life cycle hypothesis: a comment', *European Economic Review*, **33**, 1489–98.

Kotlikoff, Laurence J. and Lawrence H. Summers (1981), 'The role of intergenerational transfers in aggregate capital accumulation', *Journal of Political Economy*, **89**, August, 706–32.

Kydland, Finn E. and D'Ann M. Petersen (1997), 'Does being different matter?', Federal Reserve Bank of Dallas *Economic Review*, Third Quarter, 2–11.

Laitner, John and Thomas Juster (1996), 'New evidence on altruism: a study of TIAA-CREF retirees', *American Economic Review*, **86**, September, 893–908.

Laxton, Douglas, Peter Isard, Hamid Faruqee, Eswar Prasad and Bart Turtelboom (1998), 'MULTIMOD Mark III: the core dynamic and steady-state models', Occasional Paper No. 164, Washington, DC: International Monetary Fund, June.

Lucas, Robert E., Jr. (1967), 'Adjustment costs and the theory of supply', *Journal of Political Economy*, **75**, 321–34.

Lusardi, Annamaria (1998), 'On the importance of the precautionary saving motive', *American Economic Review, Papers and Proceedings*, **88**, May, 449–53.

Mason, Andrew (1987), 'National saving rates and population growth: a new model and new evidence', in D. Gale Johnson and Ronald D. Lee (eds), *Population Growth and Economic Development: Issues and Evidence*, Madison, WI: University of Wisconsin Press, pp. 523–70.

Mason, Andrew (1988), 'Saving, economic growth, and demographic change', *Population and Development Review*, **14**, 113–44.

Masson, Paul R. (1992), 'Effects of long-run demographic changes in a multi-country model', in Colin Hargreaves (ed.), *Macroeconomic Modelling of the Long Run*, Aldershot, UK and Brookfield, VT: Edward Elgar.

Masson, Paul R., Tamim Bayoumi and Hossein Samiei (1995), 'Saving behavior in industrial and developing countries', *Staff Studies for the World Economic Outlook*, Washington, DC: International Monetary Fund, September.

Masson, Paul R., Tamim Bayoumi and Hossein Samiei (1998), 'International evidence on the determinants of private saving', *World Bank Economic Review*, **12**, 483–501.

Masson, Paul R., Guy Meredith and Steven A. Symansky (1990), 'MULTIMOD Mark II: a revised and extended model', Occasional Paper No. 71, Washington, DC: International Monetary Fund.

Masson, Paul R. and Ralph W. Tryon (1990), 'Macroeconomic effects of projected population aging in industrial countries', *IMF Staff Papers*, **37**, September, 453–85.

McKibbin, Warwick J. (1997), 'The macroeconomic experience of Japan since 1990: an empirical investigation', Brookings Discussion Paper in International Economics No. 131, Washington, DC: Brookings Institution, June.

McKibbin, Warwick J. and Jeremy Nguyen (2001), 'The impacts of demographic change in Japan: some preliminary results from the MSG3 model', Paper presented at the Third International Forum of Collaboration Projects sponsored by the Economic and Social Research Institute, Japanese Cabinet Office, Tokyo, September.

McKibbin, Warwick J. and Jeffrey D. Sachs (1991), *Global Linkages: Macroeconomic Interdependence and Cooperation in the World Economy*, Washington, DC: Brookings Institution.

McKibbin, Warwick J. and David Vines (2000), 'Modelling reality: the need for both intertemporal optimization and stickiness in models for policy-making', *Oxford Review of Economic Policy*, **16** (4), 106–37.

McKibbin, Warwick J. and Peter J. Wilcoxen (1997), 'Macroeconomic volatility in general equilibrium', Brookings Discussion Paper in International Economics No. 140, Washington, DC: Brookings Institution.

McKibbin, Warwick J. and Peter J. Wilcoxen (1999), 'The theoretical and empirical structure of the G-cubed model', *Economic Modelling*, **16**, January, 123–48.

Menchik, Paul and Martin David (1983), 'Income distribution, lifetime saving and bequests', *American Economic Review*, **73**, September, 673–90.

Meredith, Guy (1991), 'A steady-state version of MULTIMOD', Mimeo, Washington, DC: International Monetary Fund, July.

Meredith, Guy (1995), 'Demographic changes and household saving in Japan', and 'Alternative long-run scenarios', in Ulrich Baumgartner and Guy Meredith (eds), *Saving Behavior and the Asset Price 'Bubble' in Japan*, IMF Occasional Paper No. 124, Washington DC: International Montary Fund, April.

Modigliani, Franco (1970), 'The life-cycle hypothesis of saving and intercountry differences in the saving ratio', in W.A. Eltis, M.F.G. Scott and J.N. Wolfe (eds), *Induction, Growth, and Trade: Essays in Honor of Sir Roy Harrod*, Oxford: Oxford University Press, 197–225.

Modigliani, Franco and Richard Brumberg (1954), 'Utility analysis and the consumption function: an interpretation of cross-section data', in *Post-Keynesian Economics*, New Brunswick, NJ: Rutgers University Press, pp. 388–436.

Modigliani, Franco and Richard Brumberg (1979), 'Utility analysis and the consumption function: an attempt at integration', in Andrew Abel (ed.), *The Collected Papers of Franco Modigliani*, vol. 2, Cambridge, MA: MIT Press, pp. 128–97.

Modigliani, Franco and Arlie G. Sterling (1983), 'Determinants of private saving with special reference to the role of social security: cross-country tests', in F. Modigliani and R. Hemming (eds), *The Determination of National Saving and Wealth*, proceedings of a conference held by the International Economic Association, Bergamo, Italy, 9–14 June, 1980, London: Macmillan, pp. 24–55.

Muellbauer, John and Ralph Lattimore (1995), 'The consumption function: a theoretical and empirical overview', in M. Hashem Pesaran and M.R. Wickens (eds), *Handbook of Applied Econometrics: Macroeconomics*, Oxford: Blackwell, pp. 221–311.
Organization for Economic Cooperation and Development (OECD) (1990), *OECD Economic Surveys: Japan, 1989/90*, Paris: OECD.
Paxson, Christina H. (1996), 'Saving and growth: evidence from micro data', *European Economic Review*, **40**, 255–88.
Poterba, James M. (ed.) (1994), *International Comparisons of Household Saving*, Chicago: University of Chicago Press for the National Bureau of Economic Research.
Ríos-Rull, José-Víctor (1996), 'Life-cycle economies and aggregate fluctuations', *Review of Economic Studies*, **63**, July, 465–89.
Ríos-Rull, José-Víctor (2001), 'Population changes and capital accumulation: the aging of the baby boom', manuscript, Philadelphia: University of Pennsylvania.
Shibuya, Hiroshi (1987), 'Japan's household saving rate: an application of the life cycle hypothesis', IMF Working Paper No. 87/15, Washington, DC: International Monetary Fund, March.
Storesletten, Kjetil (1995), 'Immigration policy and the aggregate savings rate', in 'The Economics of immigration', unpublished PhD thesis, Carnegie-Mellon University, Pittsburgh, PA.
Storesletten, Kjetil (2000), 'Sustaining fiscal policy through immigration', *Journal of Political Economy*, **108**, April, 300–323.
Taylor, Alan M. and Jeffrey G. Williamson (1994), 'Capital flows to the New World as an intergenerational transfer', *Journal of Political Economy*, **102**, April, 348–69.
Tobin, James (1967), 'Life cycle saving and balanced growth', in William Fellner (ed.), *Ten Studies in the Tradition of Irving Fisher*, New York: Wiley, pp. 231–56.
Treadway, A. (1969), 'On rational entrepreneurial behavior and the demand for investment', *Review of Economic Studies*, **36**, 227–39.
Turner, Dave, Claude Giorno, Alain De Serres, Ann Vourc'h and Pete Richardson (1998), 'The macroeconomic implications of ageing in a global context', OECD Economics Department Working Paper No. 193, Paris: OECD, March.
Uzawa, Hirofumi (1969), 'Time preference and the Penrose effect in a two-class model of economic growth', *Journal of Political Economy*, **77**, 628–52.
Velculescu, Delia (2000), 'Implications of habit formation for demographic change and the US saving rate', Mimeo, Johns Hopkins University, Baltimore, MD.
Weil, David N. (1994), 'The saving of the elderly in micro and macro data', *Quarterly Journal of Economics*, **109**, February, 55–81.
Weil, Philippe (1989), 'Overlapping families of infinitely-lived agents', *Journal of Public Economics*, **38**, 183–98.
Wetterstrand, W.H. (1981), 'Parametric models for life insurance mortality data: Gomperty's law over time', *Transactions of the Society of Actuaries*, **33**, 159–75.
Wilhelm, Mark O. (1996), 'Bequest behavior and the effect of heirs' earnings: testing the altruistic model of bequests', *American Economic Review*, **89**, September, 874–92.
Yaari, Menahem E. (1965), 'Uncertain lifetime, life insurance, and the theory of the consumer', *Review of Economic Studies*, **32**, April, 137–50.

Zeldes, Stephen (1989), 'Consumption and liquidity constraints: an empirical investigation', *Journal of Political Economy*, **97**, 305–46.

Zhang, Long (1996), 'Intertemporal fiscal closure rules and fiscal policy in a continuous OLG model', Doctoral Dissertation, Department of Economics, University of California, Santa Cruz.

Index