palgrave
macmillan

Intergenerational Equity and Sustainability

Edited by

John Roemer and Kotaro Suzumura

Intergenerational Equity and Sustainability

This is IEA conference volume no. 143

This page intentionally left blank

Intergenerational Equity and Sustainability

Edited by

John Roemer

Yale University, USA

and

Kotaro Suzumura

Hitotsubashi University, Japan

palgrave
macmillan

in association with
Palgrave Macmillan

Softcover reprint of the hardcover 1st edition 2007 978-0-230-00786-4

First published 2007 by
PALGRAVE MACMILLAN
Houndmills, Basingstoke, Hampshire RG21 6XS and
175 Fifth Avenue, New York, N.Y. 10010
Companies and representatives throughout the world

PALGRAVE MACMILLAN is the global academic imprint of the Palgrave Macmillan division of St. Martin's Press, LLC and of Palgrave Macmillan Ltd. Macmillan® is a registered trademark in the United States, United Kingdom and other countries. Palgrave is a registered trademark in the European Union and other countries.

ISBN 978-1-349-28344-6 ISBN 978-0-230-23676-9 (eBook)
DOI 10.1057/9780230236769

This book is printed on paper suitable for recycling and made from fully managed and sustained forest sources.

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

Library of Congress Cataloging-in-Publication Data

International Economic Association Roundtable Meeting on Intergenerational Equity (2005 : Hakone-machi (Japan))
Intergenerational equity and sustainability / edited by John Roemer and Kotaro Suzumura.
p. cm.

Papers presented at the International Economic Association Roundtable Meeting on Intergenerational Equity on March 10–12, 2005 in Hakone, Japan.
Includes bibliographical references and index.

1. Welfare economics—Congresses. 2. Intergenerational relations—Economic aspects—Congresses. 3. Distributive justice—Congresses. 4. Sustainable development—Congresses. I. Roemer, John E. II. Suzumura, Kotaro, 1944– III. Title.
HB846.I58 2005
339.2′2—dc22 2006052502

10 9 8 7 6 5 4 3 2 1
16 15 14 13 12 11 10 09 08 07

Contents

List of Figures

The International Economic Association

A non-profit organization with purely scientific aims, the International Economic Association (IEA) was founded in 1950. It is a federation of some 60 national economic associations in all parts of the world. Its basic purpose is the development of economics as an intellectual discipline, recognising a diversity of problems, systems and values in the world and taking note of methodological diversities.

The IEA has, since its creation, sought to fulfil that purpose by promoting mutual understanding among economists through the organization of scientific meetings and common research programmes, and by means of publications on problems of fundamental as well as of current importance. Deriving from its long concern to assure professional contacts between East and West and North and South, the IEA pays special attention to issues of economies in systemic transition and in the course of development. During its fifty years of existence, it has organized more than a hundred round-table conferences for specialists on topics ranging from fundamental theories to methods and tools of analysis and major problems of the present-day world. Participation in round tables is at the invitation of a specialist programme committee, but 13 triennial World Congresses have regularly attracted the participation of individual economists from all over the world.

The Association is governed by a Council, composed of representatives of all member associations, and by a 15-member Executive Committee which is elected by the Council. The Executive Committee (2002–2005) at the time of the Hakone Conference was:

President:	Professor János Kornai, Hungary
Vice-President:	Professor Bina Agarwal, India
Treasurer:	Professor Jacob Frenkel, Israel
Past President:	Professor Robert Solow, USA
President-elect:	Professor Guillermo Calvo, Argentina
Other members:	Professor Maria Augusztinovics, Hungary
	Professor Eliana Cardoso, Brazil
	Professor Duardo Engel, Chile
	Professor Heba Handoussa, Egypt
	Professor Michael Hoel, Norway
	Professor Jean-Jacques Laffont, France
	Professor Andreu Mas Colell, Spain

	Professor Kotaro Suzumura, Japan
	Professor Alessandro Vercelli, Italy
Advisers:	Professor Fiorella Kostoris Padoa Schioppa, Italy
	Professor Vitor Constancio, Portugal
Secretary-General:	Professor Jean-Paul Fitoussi, France
General Editor:	Professor Michael Kaser, UK

Sir Austin Robinson was an active Adviser on the publication of IEA Conference proceedings from 1954 until his final short illness in 1993.

The Association has also been fortunate in having secured many outstanding economists to serve as President:

Gottfried Haberler (1950–53), Howard S. Ellis (1953–56), Erik Lindahl (1956–59), E.A.G. Robinson (1959–62), Ugo Papi (1962–65), Paul A. Samuelson (1965–68). Erik Lundberg (1968–71), Fritz Machlup (1971–74), Edmond Malinvaud (1974–77), Shigeto Tsuru (1977-80), Victor L. Urquidi (1980–83), Kenneth J. Arrow (1983–86), Amartya Sen (1986–89), Anthony B. Atkinson (1989–1992), Michael Bruno (1992–95), Jacques Drèze (1995–99), Robert M. Solow (1999–2002) and Janos Kornai (2002–05).

The activities of the Association are mainly funded from the subscriptions of members and grants from a number of organizations. Support from UNESCO since the Association was founded, and from its International Social Science Council, is gratefully acknowledged.

Acknowledgements

The Conference was held on 10 to 12 March 2005 at the Hotel de Yama, Hakone, Japan. The responsibility for the selection of papers was in the hands of a Programme Committee with the following members:

Walter Bossert
John Roemer (co-chair)
Joaquim Silvestre
Kotaro Suzumura (co-chair)
Koichi Tadenuma.

To cover a wide spectrum of issues related to intergenerational equity and sustainability, the Committee invited 25 highly-qualified economists and philosophers to discuss 17 papers over three full days. The IEA and the Chairman express sincere gratitude to all who contributed, as committee members, authors and discussants, without whose collaboration this book would not have been fruitfully completed.

An international academic meeting of this size and level could not have been organized without the generous support of a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan. Warm thanks are due to those involved in the planning and implementation of the Round Table and in particular to the 21st Century Centre of Excellence Programme on the Normative Evaluation and Social Choice of Contemporary Economic Systems, and the Project on Intergenerational Equity, both at Hitotsubashi University, and to the Union of National Economic Associations in Japan for supplementary financial help.

Last but not least, thanks are also due to the staff of the 21st Century Centre of Excellence Programme and the Project on Intergenerational Equity, particularly Takane Fukushima, Kana Ito, Yuko Kajiyama, Ayako Komoto, Aki Soejima, Chieko Takada and Harumi Uwai, whose dedicated work were instrumental in bringing the meeting to a successful conclusion.

The IEA General Editor, Michael Kaser, was responsible for the editing.

List of Contributors

Geir Asheim, University of Oslo, Norway
Claude d'Aspremont, Université de Louvain, Belgium
Kaushik Basu, Cornell University, USA
Charles Blackorby, University of Warwick, UK
Walter Bossert, Université de Montréal, Canada
David Donaldson, University of British Columbia, Canada
Marc Fleurbaey, Université de Pau et des Pays de l'Adour, France
Toshihiro Ihori, University of Tokyo, Japan
Tapan Mitra, Cornell University, USA
John Roemer, Yale University, USA
Tomoichi Shinotsuka, University of Tsukuba, Japan
Joaquim Silvestre, University of California at Davis, USA
Koichi Suga, Waseda University, Japan
Kotaro Suzumura, Hitotsubashi University, Japan
Koichi Tadenuma, Hitotsubashi University, Japan
Noriyuki Takayama, Hitotsubashi University, Japan
Bertil Tungodden, Norwegian School of Economics and Business Administration, Bergen, Norway
Peter Vallentyne, University of Missouri at Columbia, USA
Ngo Van Long, McGill University, Canada
Roberto Veneziani, Queen Mary, University of London, UK
Yongsheng Xu, Georgia State University, USA
Naoki Yoshihara, Hitotsubashi University, Japan

List of Abbreviations and Acronyms

APE	anonymous Paretian egalitarianism
C	continuity
CES	constant elasticity of substitution
CIREQ	Centre Universitaire de recherche en économique quantitative, Université de Montreal
CLS	cosmopolitan leximin solution
CPI	consumer price index
EG	Egalitarian Equivalence
EL	excess liabilities
ELLR	Equity in Lifetime Rate of Return
EO	equal opportunity
EWEP	Equal Welfare for Equal Preferences
EWUP	Equal Welfare for Uniform Preferences
FOC	first order condition
FP	fixed population
FQRSC	Fonds québécois de la recherche sur la société et la culture
FRP	fixed relative position
GC	generational continuity
GDP	gross domestic product
GP	guaranteed pension
GPI	generational Pareto indifference
HDI	Human Development Index
HE	Hammond Equity
HEF	Hammond Equity for the Future
IF	Independent Future
J-HD	J-Reference Human Development
KNH	Kosei-Nenkin-Hoken
Leximin	lexicographic completion of the maximin
LNBS	leximin Nash bargaining solution
NBER	National Bureau of Economic Research
NCLS	non-cooperative leximin solution
NDG	notional defined contribution
NE	No-Envy
NEEG	No-Envy among Equally-Endowed Generations
NELC	No-Envy in Lifetime Consumptions
NEOC	No-Envy in Overlapping Consumptions
NEUG	No-Envy among Uniform-Endowed Generations

NIPSSR	National Institute of Population and Social Security Research
NSPE	non-selfish Pareto efficiency
O	order
OLG	overlapping generations
PA	person-affecting
PB	public bads monotonicity
PE	Pareto efficiency
QT	quasi-transitivity
R&D	research & development
RC	restricted continuity
RFG	Responsibility for Future Generations
RS	restricted sensitivity
RUSC	restricted upper semi-continuity
RWP	restricted weak Pareto
SOF	social ordering function
SP	strong Pareto
SQOF	social quasi-ordering function
SS	strong sensitivity
SSHRC	Social Sciences and Humanities Research Council (Canada)
STP	sensitivity to the present
SWF	social welfare function(al)
SWO	social welfare ordering
SWQ	social welfare quasi-ordering
TFR	total fertility rate
UNEP	Undomination among Equal Preferences
USC	upper semi-continuity
WD	weak dominance
WLD	weak Lorenz domination
WP	weak Pareto
WPD	weak Pigou-Dalton
WS	weak sensitivity

Introduction

John Roemer and Kotaro Suzumura

1 Historical background

The problems of intergenerational equity, efficiency, and rationality have been receiving close attention since the advent of modern normative economics. To start with, there is a strong utilitarian tradition of treating otherwise equal generations equally. Henry Sidgwick expounded this tradition lucidly as follows: '[I]t may be asked, How far we are to consider the interests of posterity when they seem to conflict with those of existing human beings? It seems... clear that the time at which a man exists cannot affect the value of his happiness from a universal point of view; and that the interests of posterity must concern a Utilitarian as much as those of his contemporaries' (Sidgwick 1907, p. 414). To paraphrase this viewpoint in the modern parlance, the intergenerational normative judgements should satisfy the requirement of finite anonymity with respect to the date at which an individual is born.

Sidgwick's view was endorsed by Arthur Pigou, Frank Ramsey, and Roy Harrod to cite just a few salient scholars who followed him. For Pigou, who laid the foundation of the so-called 'old' welfare economics, disobeying the anonymity principle *à la* Sidgwick is nothing other than being irrational: '[E]verybody prefers present pleasures or satisfactions of given magnitude to future pleasures or satisfactions of equal magnitude, even when the latter are perfectly certain to occur. ... [T]his preference for present pleasures... implies only that our telescopic faculty is defective, and that we... see future pleasures, as it were, on a diminished scale' [Pigou (1920, pp. 24–25)]. Ramsey, who pioneered the theory of optimal saving, went so far as to say that a practice of discounting later enjoyments vis-à-vis earlier ones was 'ethically indefensible,' a practice which arises 'merely from the weakness of the imagination [Ramsey (1928, p. 543)].' Harrod, who initiated the postwar growth theory, went even further, and asserted that '[w]e do not see the future so vividly as the present and underrate the advantage of having money at a future date compared with that of having it now... Also we may be dead at the future date and not rate the welfare of our heirs as highly as our own.

The desire to use the money now is reinforced by animal appetite. Greed may be thought to be as appropriate a name for this attitude as time preference, though less dignified. Time preference in this sense is a human infirmity, probably stronger in primitive than in a civilized man [Harrod (1948, p. 37)].' Thus, the widespread convention of discounting future utilities or welfares vis-à-vis present utilities or welfares is viewed by these celebrated scholars as ethically indefensible, reflecting human irrationality and fallibility.

Tjalling Koopmans (1960) raised a strong criticism against this view by showing that the rational, continuous, and stationary evaluation of infinite resource allocation programmes cannot but exhibit a phenomenon which he christened *impatience*, viz., the preference for an advancement along the time axis of an outcome yielding higher utility vis-à-vis another outcome yielding lower utility. This intriguing proposition was further elaborated by Peter Diamond (1965) into a general impossibility theorem to the effect that there exists no social evaluation ordering over the set of infinite utility streams satisfying the following three basic requirements: the *Pareto efficiency principle*, the *equity principle à la Sidgwick*, and the *continuity axiom* on the social evaluation ordering with respect to the sup topology. It is no wonder that this impossibility theorem caused a stir in the profession, as it means that there is a fundamental conflict between the intergenerational equity principle à la Sidgwick and the widely accepted Pareto efficiency principle in the presence of a weak regularity requirement of continuity on the social evaluation ordering. Confronting this impossibility theorem is one way that we may deepen our understanding of the relationship among intergenerational equity, efficiency, and rationality, but the conceptual framework of Koopmans and Diamond may not be sufficiently articulated to accommodate the variegated facets of the basic problem.

Recall that the Koopmans-Diamond framework expressed the problem of intergenerational equity in terms of the social evaluation ordering over the set of infinite utility streams, where a generic component u_t of representative stream $\boldsymbol{u} = (u_1, u_2, \ldots, u_t, \ldots)$ stands for the utility enjoyed by generation t. A salient feature of this description is that there is neither an explicit channel through which different generations interact with each other, nor any explicit stipulation of resource constraints. As in the case of the pure theory of consumer behaviour, where a consumer's preference ordering over the set of all conceivable consumption programmes is described prior to the introduction of the budget constraint that reflects the market environment constraining the consumer, the Koopmans-Diamond framework may well qualify as a convenient beginning for the theory of intergenerational equity. However, it falls short of capturing many other intergenerational issues of crucial importance. To see this, we cite three concrete problems of intergenerational equity, efficiency, and rationality.

In the first place, the problem of designing and implementing an equitable and efficient mechanism for intergenerational transfer of resources seems to call for a different conceptual framework altogether. The required framework is one where the resource transfer between adjacent generations is explicitly formulated within the model. An ingenious model, most helpful in this context, was introduced by Paul Samuelson (1958), widely known as the overlapping generations model. Most, if not all, attempts to capture the essence of social security schemes for intergenerational transfers are based on a suitably augmented version of the Samuelsonian model. Note, however, that the traditional focus of these analyses including Samuelson's own has been on the intertemporal efficiency, and not on the intergenerational equity, of resulting resource allocations over generations. This is a conspicuous lacuna in the literature. In trying to work on this open question, our analytical focus should be on the balance between overlapping generations as to their claims for equitable treatment without sacrificing the intertemporally efficient use of scarce resources.

In the second place, consider the concept of *sustainable development*. This phrase was first introduced in 1980 by the International Union for the Conservation of Nature and Natural Resources. It came to acquire worldwide recognition through the publication in 1987 of the so-called Brundtland Report by the World Commission on Environment and Development. According to this report, '[s]ustainable development is development that meets the needs of the present without compromising the ability of future generations to meet their own needs [World Commission on Environment and Development (1987, p.43)].' It is worthwhile to emphasize that the proposed concept of sustainable development embodies a basic belief to the effect that the well-being of future generations is as important as the well-being of the present generation. In this sense, sustainability echoes the classical principle of Sidgwick. According to this basic principle, all generations, irrespective of when they emerged in the past or will emerge in the future, should have equal opportunity to lead worthwhile lives. In order to give concrete substance to this admittedly abstract concept, we must go beyond the narrow informational framework of Koopmans and Diamond, and take account, in our analysis, of the limitations imposed by the past, present and future states of technology.

In the third place, consider the long-run problem of environmental externalities, viz., the issue of global warming. Economic activity since the Industrial Revolution has contributed to the accumulation of greenhouse gases, which will affect the well-being of future generations. Thus, global warming poses a unique externality problem in which many of those who should be held responsible no longer exist, and many of those who will be most severely affected have not yet been born. As if this is not enough, there is yet another intriguing feature of the problem of future generations in the context of global warming. Observe that it is undeniable that the actual historical path of human evolution up to the present is determined by the choices made by past generations, and the population of all generations up to the present generation is also determined. Indeed, human evolution in the future hinges on choices made by the present generation. For instance, what our generation does with respect to energy consumption cannot but affect the size and personality characteristics of the next generation. The malleability of future generations in this sense is an aspect of the non-identity problem posed by Derek Parfit (1984). This fact makes the problem of intergenerational equity, efficiency, and rationality in the presence of global warming even more complex than otherwise.

Thus, the conceptual framework of the Koopmans–Diamond model is too thin to accommodate many issues that we face. More highly articulated models, and an informational framework that goes beyond welfarism, or even consequentialism, may well be needed to provide a satisfactory analysis of the problem.

2 Brief overview of the book

Based on the papers presented at the Hakone Roundtable Meeting, the book consists of five parts, each containing three or four chapters. In this section we briefly describe the contents of the book.

Part I (Equity among Overlapping Generations) contains three chapters which focus on intergenerational equity and efficiency in overlapping generations economies. Chapter 1 by Toshihiro Ihori ('Pension Contributions and Capital Accumulation') investigates some dynamic implications of pension contributions and intergenerational transfers under the pay-as-you-go scheme, where each interest group may easily avoid contributions and the national government cannot effectively penalize such evasive behaviour. Within the standard model of overlapping generations economies that incorporates subsidies from the national government, some long-run properties of public spending, social security contributions, and economic growth are explored. It is shown, among other things, that the social security contribution is too little in terms of static efficiency, but it may be too much or too little in terms of dynamic efficiency. Chapter 2 by Tomoichi Shinotsuka, Koichi Suga, Kotaro Suzumura and Koichi Tadenuma ('Equity and Efficiency in Overlapping Generations Economies') examines several distinct notions of intergenerational equity as no-envy within the model of overlapping generations economies à la Samuelson. The notion of no-envy in overlapping consumptions stipulates that, for each time period, no person should prefer the bundle of any other person, who lives in the same period, to his own. The notion of no-envy in lifetime consumptions requires that no person should prefer the lifetime consumption plan of any other person to his own. Finally, the notion of equity in lifetime rate of return requires that the lifetime rate of return in the sense of David Cass and Menaham Yaari (1966) be equal for all generations. For each of these notions of intergenerational equity, the chapter characterizes allocations satisfying that notion and clarifies logical relations among these notions. It also examines the existence of a stationary allocation that attains maximal utility under no-envy in lifetime consumption. Chapter 3 by Noriyuki Takayama ('Social Security Pensions and Intergenerational Equity: The Japanese Case') is unique in the book, as it addresses the issue of intergenerational equity in the concrete context of a specific nation, viz., Japan. It is worth recalling that Japan has the oldest population in the world; although it has a generous social security pension programme, the income statement of its principal pension programme has been in deficit since 2001, and the balance sheet of its pension programme has been suffering from huge excess liabilities; as a result, there is a growing scepticism among the citizenry concerning the government's commitment to the public pension scheme. With this background in mind, Takayama addresses himself to intergenerational issues in the context of Japan's social security pension reform with special focus, on the one hand, on the remedy for the mistakes made in the past, and, on the other hand, on the pension scheme which would avoid intergenerational inequities arising from uncertainties in the future.

Part II (Ranking Infinite Utility Streams) turns to the classical Sidgwickian concern with intergenerational equity, efficiency, and rationality in the context of evaluating infinite utility streams. Chapter 4 by Geir Asheim, Tapan Mitra and Bertil Tungodden ('A New Condition for Infinite Utility Streams and the Possibility of Being Paretian') introduces a new equity condition called Hammond Equity for the Future that replaces the Sidgwick equity condition. The new condition captures the ethical idea that a sacrifice by the present generation leading to a uniform gain for all future generations cannot bring about a less desirable utility stream if the present generation remains better off than future generations even after making the sacrifice. The effect of replacing the Sidgwick equity condition with Hammond Equity for the Future in the presence of other conditions *à la* Diamond is carefully examined, and leads to another impossibility theorem. Chapter 5 by Kaushik Basu and Tapan Mitra ('Possibility Theorems for Equitably Aggregating Infinite Utility Streams') tries to identify an escape route from the Diamond impossibility theorem by seeking the existence of a real-valued social welfare function over the set of infinite utility streams, which is not necessarily continuous, by weakening the Pareto efficiency principle and exploring some domain restrictions. Chapter 6 by Tapan Mitra and Kaushik Basu ('On the Existence of Paretian Social Welfare Quasi-Orderings for Infinite Utility Streams with Extended Anonymity') focuses on the generalizability of the finite anonymity axiom in the context of a social welfare quasi-ordering on the set of infinite utility streams and in the presence of the Pareto principle. The Sidgwick–Pigou–Ramsey equity axiom is usually captured by a finite anonymity axiom saying that two infinite utility streams, one of which can be obtained from the other by applying a finite permutation, should be declared socially indifferent. The attempt to expand the admissible class of permutations beyond that of finite permutations has a clear limit in that arbitrary infinite permutations can be easily checked to be in conflict with the Pareto principle. This chapter is devoted to the logical problem of characterizing the class of admissible permutations which makes the resulting anonymity axiom compatible with the Pareto principle in the presence of social welfare quasi-orderings. Chapter 7 by Yongsheng Xu ('Pareto Principle and Intergenerational Equity: Immediate Impatience, Universal Indifference and Impossibility') is designed to answer two questions. Its first purpose is to examine the scope for obtaining a possibility theorem for the social evaluation relation to be Paretian and intergenerationally equitable. Its second purpose is to crystallize the respective implications of the Paretian axiom and the intergenerational equity axiom for a social welfare function, thereby clarifying the structure of a social evaluation relation satisfying these two axioms.

Part III (Intergenerational Evaluations) is still concerned with the social evaluation of infinite utility streams, but it gathers contributions which go beyond the classical Koopmans–Diamond conceptual framework. Chapter 8 by Claude d'Aspremont ('Formal Welfarism and Intergenerational Equity') looks closely at the foundations of the Koopmans–Diamond formulation within the social welfare functional approach to social choice theory, which is due to Amartya Sen (1970). Recall that the classical framework is welfarist in nature in that utilities of all generations exhaust all the information required for the construction of a social evaluation on the future history of infinite length. Besides, the classical framework also presupposes that, within each generation, the individual utilities are aggregated into a single generational utility. The chapter discusses some disturbing implications of these two presumptions of the traditional approach. Chapter 9 by Charles Blackorby, Walter Bossert and David Donaldson ('Intertemporal Social Evaluation') employs an axiomatic approach to identify ethically attractive social evaluation procedures for intertemporal social evaluation. In particular, they explore the possibilities of using welfare information as well as non-welfare information in a model of intertemporal social evaluation. For the relevant non-welfare information, they adopt the birth dates and lengths of the individuals' lives, whereas they adopt lifetime utilities as the relevant welfare information. They identify several sets of axioms which characterize the birth-date dependent or lifetime-dependent versions of generalized utilitarianism. Chapter 10 by Marc Fleurbaey ('Intergenerational Fairness') studies the construction of a criterion for the ethical evaluation of allocations in an overlapping generations economy with linear technology and heterogeneous preferences. It studies how to construct a ranking of allocations on the basis of axioms of efficiency and fairness, and characterizes a particular ranking of the infimum kind, which evaluates every individual situation in terms of money metric utility. It also shows that defining a complete social ordering satisfying the efficiency and fairness requirements raises a difficulty due to the infinite horizon of the basic model. Chapter 11 by Bertil Tungodden and Peter Vallentyne ('Person-Affecting Paretian Egalitarianism with Variable Population Size') develops and discusses a deontic version of anonymous Paretian egalitarianism for the variable population case where who comes into existence depends on agents' choices. In more specific terms, this is done in the context of a person-affecting framework where an option is just if and only if it wrongs no one according to certain plausible conditions on what constitutes wrong-doing.

The chapters in Part IV model sustainability as a path of intertemporal consumption that guarantees a lower bound on the utility of all generations. The natural goal is to maximize the size of this bound – thus, maximin (or leximin) utility of all generations that will exist. Roemer's chapter ('Intergenerational Justice and Sustainability under the Leximin Ethic') postulates a renewable resource (the 'forest' or 'biosphere') which can be 'harvested' by each generation to produce a consumption good, and which is enjoyed as well in its natural state. The intergenerational leximin solution is defined and studied with exogenous technical progress; in particular, conditions are given under which the leximin solution is compatible with increasing utility of generations over time. Finally, there is some study of endogenous technical progress. Roemer and Veneziani ('Intergenerational Justice, International Relations and Sustainability') study a world with two representative agents at each date, one in the poor South and the other in the rich North. There is a global commons (the biosphere) which each agent pollutes in order to produce a consumption good; the technology of the South is inferior, so it must pollute more per unit of output produced. The consequences of three regimes of international cooperation (or lack of it) are studied, ranging from non-cooperative Nash equilibrium between the South and the North, to the 'cosmopolitan' solution, where national borders are ignored. The analysis points out the conflicts between environmental concerns and sustainability, and between international and intergenerational justice. Silvestre's chapter ('Intergenerational Equity and Human Development') applies two normative criteria, viz., leximin and sufficientarianism (which requires a 'good enough' standard of living), to an overlapping-generations society with a given decision date. These criteria may conflict with the desideratum of human development across generations, because they advocate benefitting the early generations, presumably the worst-off ones, assuming that technological progress occurs. Two difficulties in the application of the leximin criterion are discussed. First, leximin requires the exogenous specification of the initial obligations to decision-date seniors. Second, the appearance of human development at the leximin solution imposes an extreme lifetime pattern, with low levels of primary goods during the productive years, and abundance at the late stages of life. Sufficientarianism attenuates these difficulties, and is to some extent complementary to leximin when applied to the intergenerational problem.

Part V (Long-Run Issues of Intergenerational Equity) is assigned to three chapters which are devoted to the long-run issues of intergenerational equity and justice. Chapter 15 by Ngo Van Long ('Toward a Just Savings Principle') starts from a brief review of John Rawls' (1971) objection to the direct application of his difference principle, which is the core concept of his principles of justice, to the problem of intergenerational equity, viz., the problem of just savings. It then makes some adjustments to the difference principle so as to crystallize the author's proposed concept of just savings, and extracts some implications of these adjustments by means of some formal economic models. Chapter 16 by Kotaro Suzumura and Koichi Tadenuma ('Normative Approaches to the Issues of Global Warming: Responsibility and Compensation') focuses on the analysis of equitable treatment among far distant generations, a prototypical example thereof being the issue of global warming. This is a severe problem of environmental externalities. Unlike many other problems of environmental disruptions, this problem has a unique feature: a large part of those who are responsible for causing the problem do not exist anymore, and a large part of those who will be most severely affected by the problem do not yet exist. This crucial feature prevents standard resolution schemes for environmental externalities from being applied to this class of problems. Besides, the informational basis of normative analysis cannot but go beyond the confinement of welfarism in view of the crucial problem of malleability of future generations, which is an example of Derek Parfit's (1984) non-identity problem. In view of these novel features, the chapter proposes a new principle of responsibility and compensation in the context of global warming. Finally, Chapter 17 by Naoki Yoshihara ('Fundamental Incompatibility among Economic Efficiency, Intergenerational Equity and Sustainability') examines intergenerational resource allocations in production economies with long-run negative externalities and shows that the basic axioms of economic efficiency, intergenerational equity and environmental sustainability cannot but conflict with each other.

Since the issues of intergenerational equity, efficiency, and rationality are broad, and there are many alternative approaches which we may conceive in trying to do justice to these issues, it is too much to hope that the Hakone Roundtable Meeting and the present book based on our extensive discussions at the Meeting could settle these perennial issues. Indeed, further studies, which are meant to pursue some of the issues raised but left open at the Hakone Meeting, are being vigorously promoted by some of the participants of the Hakone Meeting. It is hoped that this book will serve as a useful reference book as well as an indispensable bridge between what precedes it and what succeeds it on the issues of intergenerational equity, efficiency and rationality.

References

Anand, S. and A. K. Sen (1996) 'Sustainable Human Development: Concepts and Properties', United Nations Development Programme, Office of Development Studies.

Arrow, K. J. (1951) *Social Choice and Individual Values* (New York; Wiley: 2nd edn, 1963).

Broome, J. (1992) *Counting the Cost of Global Warming* (Cambridge: White House Press).

Cass, D. and M. Yaari (1966) 'A Re-examination of the Pure Consumption Loans Model', *Journal of Political Economy*, vol. 74, pp. 353–67.

Dasgupta, P. (2001) *Human Well-Being and the Natural Environment* (Oxford: Oxford University Press).

Diamond, P. (1965) 'The Evaluation of Infinite Utility Streams,' *Econometrica*, Vol. 33, pp. 170–7.

Dworkin, R. (2000) *Sovereign Virtue: The Theory and Practice of Equality* (Cambridge, MA: Harvard University Press).

Harrod, R. (1948) *Towards a Dynamic Economics* (London: Macmillan).

Heal, G. (1998) *Valuing the Future: Economic Theory and Sustainability* (New York: Columbia University Press).

Koopmans, T. (1960) 'Stationary Ordinal Utility and Impatience', *Econometrica*, vol. 28, pp. 287–309.

Parfit, D. (1984) *Reasons and Persons* (Oxford: Oxford University Press).

Pigou, A. C. (1920) *The Economics of Welfare* (London: Macmillan; 4th edition, 1932).

Ramsey, F. P. (1928) 'A Mathematical Theory of Saving', *Economic Journal*, vol. 38, pp. 543–59.

Rawls, J. (1971) *A Theory of Justice* (Cambridge, MA: Harvard University Press).

Samuelson, P. A. (1958) 'An Exact Consumption-Loan Model of Interest With or Without the Social Contrivance of Money', *Journal of Political Economy*, vol. 66, pp. 467–82.

Sen, A. K. (1970) *Collective Choice and Social Welfare* (San Francisco: Holden-Day; Republished, Amsterdam: North-Holland, 1979).

Sen, A. K. (1985) *Commodities and Capabilities* (Amsterdam: North-Holland).

Sidgwick, H. (1907) *The Methods of Ethics*, 7th edition (London: Macmillan).

Suzumura, K. (1999) 'Consequences, Opportunities, and Procedures', *Social Choice and Welfare*, vol. 16, pp. 17–40.

Suzumura, K. (2000) 'Welfare Economics Beyond Welfarist-Consequentialism', *Japanese Economic Review*, vol. 51, pp. 1–32.

World Commission on Environment and Development (1987) *Our Common Future* (Oxford: Oxford University Press).

Part I
Equity Among Overlapping Generations

1
Pension Contributions and Capital Accumulation*

Toshihiro Ihori

1.1 Introduction

In recent times, much attention has been given to the long-run macroeconomic and intergenerational redistribution effects of public pension reform. It is well recognized that the pay-as-you-go system is not attractive when the rate of population growth is declining in an ageing society. However, the movement from pay-as-you-go financing to full funding is hard in terms of intergenerational equity, just as reducing the public debt–GDP ratio is hard. Researchers have investigated mechanisms under which a decentralized economy might successfully change from a public pay-as-you-go pension scheme to private fully funded schemes. There have been several important attempts to investigate such pension reforms. The standard analysis is of simulation studies using overlapping generations models based on Auerbach and Kotlikoff (1987). Cifuentes and Valdes-Prieto (1997), among others, offer the insightful result from a simulation model that describes the transition in detail, year by year. Mulligan and Sala-i-Martin (1999a,b) present a useful survey of various theories of social security.[1] Hatta and Oguchi (1999) develop the simulation study on Japanese pension reforms; see also Ihori (2002) and Oshio (2004).

Because most of the social security contributions from current workers go directly to fund benefits for current retirees in the unfunded system, the social security system does not significantly increase the level of government savings. Therefore, to the extent that social security reduces private saving, it will also tend to reduce the total level of saving in the economy. However, time-series and cross-country estimates are inconsistent and fraught with conceptual difficulties. They offer little additional information about the relationship between social security and saving. This is partly because the

* I would like to thank Marc Fleurbaey and other participants in the IEA Conference for useful comments.

actual pension system may not be regarded as pure pay-as-you-go financing, especially in Japan.

It is true that pay-as-you-go and fully funded systems are two representative pension schemes, and hence it is useful to investigate and compare the long-run implications of both systems in terms of intergenerational equity. However, the actual public pension system is not exactly the pay-as-you-go system in some countries. In particular, an important feature of Japan's system is that Japan's government has used the general fiscal account to subsidize the social security account. (For example, one-third of the basic pension benefits are subsidized from the general fiscal account; see Takayama, 1998.) One reason why Japan's government has a huge deficit is that its subsidies to the social security account are large. There is no direct link between contributions and benefits in the pension account, unlike the standard unfunded system. Such a subsidy may be regarded as an intergenerational transfer, as it allows the elderly generation to spend more than they contributed at the younger age and it is always easy to get political support for such a subsidy.

In the analysis of the long-run pension system the formulation of two relevant variables, social security benefits and contributions, would crucially affect the result. Actually, the benefit may well be regarded as exogenous and fixed for a long time as the defined benefit system. For example, in Japan the replacement ratio has been maintained at about 65 per cent since the mid-1970s. On the contrary, it seems more plausible to assume that contributions are endogenous. The Japanese government has adjusted old-age benefits and contributions based on the newest estimation of future population changes. There is little room to revise (or reduce) the replacement ratio due to the political pressure of existing old-age generations.

In this chapter we assume that the old-age benefit in real terms, which is represented by the replacement ratio, is a policy variable and hence exogenously given, while social security contributions are effectively determined by interest groups. Thus, the main political issue is to what degree the contributions should be raised.[2] For example, in Japan labour unions and firms' managers are reluctant to accept an increase in the rate of contributions. They exert political pressures to avoid a large increases in the contribution rate. Many self-employed people are simply not contributing to the basic pension. The contribution avoidance behaviour is illegal, but the pension authority does not effectively penalize it. Still in reality the pension contribution increases much even if they can easily avoid pension contributions under such poor enforcement circumstances.

It seems that all relevant interest groups (such as employees in the private sector, civil servants in the central government, local government employees, the self-employed, firms' managers and so on) may agree with an increase in contributions. Namely, when facing fiscal crises of public sector (or a cut in public goods), every interest group generally agrees with an overall increase in contributions since it would alleviate the fiscal crisis by reducing subsidies from the government's general account. But this does not necessarily imply that each interest group is willing to accept increases in its own contributions or cuts in its own privilege within the pension system. They would not readily agree with the allocation of increases in total contributions. This phenomenon may be analyzed using a concept of non-cooperative Nash equilibrium.

This chapter offers a theoretical examination of both intragenerational and intergenerational conflicts under the pay-as-you-go system with fiscal subsidies which are provided by the central government. By incorporating 'voluntary' contributions of interest groups to social security at Nash conjectures, our model can exhibit dynamic properties of pension contribution and capital accumulation. The larger the concern for public spending, the deeper accumulation of capital and larger pension contribution would likely occur. Capital accumulation and pension contributions usually grow at the same time. An increase in benefits (or an increase in the replacement ratio) will reduce the level of capital accumulation. However, a good rate of return of providing the contribution will not necessarily lead to large contributions at the second best solution. The pension contribution is too little in terms of the static efficiency (or compared with private consumption) but may be too much or too little in terms of the dynamic efficiency (or as the steady state level). Consumption taxes or combination of a subsidy to social security contributions and an interest income tax would always be desirable. If the government can control the replacement ratio, it would attain the dynamic efficiency by realizing the modified golden rule. The larger the concern for future generations, the more desirable it is to reduce the replacement ratio.

Section 1.2 presents the analytical framework of overlapping generations. Section 1.3 investigates dynamic properties of the model using a Cobb-Douglas example. Section 1.4 considers long-run implications of changes in some policy variables. Section 1.5 examines some normative aspects of public pension policy by deriving optimal taxes on consumption or subsidies on pension contributions. Finally section 1.6 offers some conclusions.

1.2 Model

Analytical framework

We develop a standard model of two-period overlapping generations. An agent (or an interest group) i of generation t born at time t, considers him/herself young in period t, old in period $t+1$, and dies at time $t+2$. When young an agent of generation t supplies one unit of labour inelastically and receives wages w_t out of which the agent consumes c^1_{it}, provides social security contributions g_{it}, pays wage taxes at the rate τ, and saves s_{it}. An agent receives capital income $(1+r_{t+1})s_{it}$ and pension benefits βw_{t+1} when retired, which the agent then spends entirely on consumption c^2_{it+1}. There are no private bequests. r_t is the rate of interest in period t. There is no population growth and each generation has n identical individuals (or interest groups). Hence, the population growth rate is always less than the real rate of return on private savings, and the pay-as-you-go system provides a lower rate of return than the funded system, which is relevant for Japan's ageing situation. Thus, younger generations would not like to pay pension contributions unless it provides some additional merits apart from pension benefits.

A member i of generation t faces the following budget constraints:

$$c^1_{it} = w_{it} - g_{it} - s_{it} - \tau w_{it} \tag{1.1}$$

$$c^2_{it+1} = (1 + r_{t+1})s_{it} + \beta w_{t+1} \tag{1.2}$$

where βw is old age benefits. β is the replacement ratio, old age benefits per average current wage income. The target contribution level is set by the government, but each interest group can reduce its own contribution by conducting political activities, as explained in section 1.1. Thus, g_i may be regarded as voluntary provision of social security contributions.

His/her lifetime utility function is written as

$$U^i_t = U^i(c^1_{it}, c^2_{it+1}, G_t) \tag{1.3}$$

where G is benefits from public spending, which is a pure public good and beneficial only for the younger aged.

The budget constraint of pay-as-you-go social security account is written as

$$n\beta w_t = z_t + \sum_{i=1}^{n} g_{it} \tag{1.4}$$

where z measures the subsidy from the government general account. The government budget constraint of this account is given by

$$qG_t + z_t = n\tau w_t \tag{1.5}$$

where q is the marginal (and average) cost of providing the public good. Equations (1.4) and (1.5) summarize the budget constraints including the social security account with fiscal subsidies from the general government account. Due to the subsidy mechanism, there is no direct link between the contributions of the young and old-age benefits in the public pension scheme, in contrast to the standard pay-as-you-go system. A subsidy from the general account of the central government makes the public pension budget constraint 'soft'. In Japan national tax revenues are used for providing basic pension benefits to the current generation.

Although there is no link between contributions and benefits in the above system, each interest group has an incentive to pay contributions since they receive the benefits from public spending, G. By introducing the subsidy mechanism and the benefit of public goods into the pay-as-you-go system, we may explain why the younger generation would like to contribute to public pensions. This is a realistic compromise for the government to collect pension contributions under the situation where it is difficult to penalize the avoidance behaviour of pension contributions. In Japan the government cannot easily collect pension contributions without agreement from the interest groups, as explained in section 1.1. For example, self-employed people could easily avoid pension contributions to the basic pension. Ihori and Itaya (2002, 2003) investigate a similar non-cooperative behaviour of interest groups using the infinite-horizon dynamic game where 'voluntary' contributions alleviate fiscal deficits in the fiscal reconstruction model.

Substituting (1.5) into (1.4), we may derive the overall budget constraint of the public sector as

$$qG_t + n\beta w_t = \tau n w_t + \sum_{i=1}^{n} g_{it} \tag{1.6}$$

which implies that an increase in pension contributions would result in an increase in public spending so long as the tax and replacement rates are fixed. In this sense, q may be regarded as an index of the (private) rate of return on pension contributions. If q is low, the contribution is very efficient in providing the public good.

From (1.1), (1.2) and (1.6), the lifetime private budget constraint is given by

$$c_{it}^{1} + \frac{1}{1+r_{t+1}} c_{it+1}^{2} + qG_t = w_t + \frac{1}{1+r_{t+1}} \beta w_{t+1} - n\beta w_t + (n-1)\tau w_t + \sum_{i \neq j} g_{it} \tag{1.7}$$

As in the standard model of voluntary provision of a pure public good, we will exclude binding contracts or cooperative behaviour between the agents and will explore the outcome of non-cooperative Nash behaviour.[3]

In this Cournot-Nash model, the right-hand side of (1.7) means 'real' income, $E_t^i \equiv w_t(1 - n\beta + (n-1)\tau) + \frac{1}{1+r_{t+1}} \beta w_{t+1} + \sum_{i \neq j} g_{it}$, which contains actual disposable income including current old age benefits, $w_t + \frac{1}{1+r_{t+1}} \beta w_{t+1}$, plus the externalities from other agents' provision of pension contributions and taxes, $\sum_{i \neq j} g_{it} + (n-1)\tau w_t$ minus total old age benefits, $n\beta w_t$.

From (1.6) and (1.7) we have

$$\sum_{i=1}^{n} E_t^i = n(1-\beta) w_t + \frac{1}{1+r_{t+1}} n\beta w_{t+1} + q(n-1)G_t \tag{1.8}$$

Namely, add (1.7) from $i=1$ to n and use (1.6). Then we will get (1.8). $q(n-1)G_t$, the third term in the right-hand side of (1.8), captures externalities from $n-1$ other persons' contributions within the same generation.

Let us then formulate the aggregate production function. The firms are perfectly competitive profit maximizers who produce output using the production function.

$$Y_t = F(K_t, n_t) = K_t^{1-\lambda} n^{\lambda} \qquad (0 < \lambda < 1) \tag{1.9}$$

As for the standard first-order conditions from the firm's maximization problem in period t, we have

$$r_t = r(K_t) \tag{1.10}$$

$$w_t = w(K_t) \tag{1.11}$$

since n is exogenously given. Since we follow the standard Diamond-type overlapping generations growth model with productive capital, capital does not depreciate at all.

In an equilibrium agents can save by holding physical capital. We have

$$ns_t = K_{t+1} \tag{1.12}$$

The system may be summarized by these two equations.

$$nE[Q(K_{t+1}), q, U_t] = n(1-\beta)w(K_t) + Q(K_{t+1})n\beta w(K_{t+1}) + q(n-1)E_2[Q(K_t), q, U_t] \tag{1.13}$$

$$nE_1[Q(K_{t+1}), q, U_t] = K_{t+1}/Q(K_{t+1}) + n\beta w(K_{t+1}) \tag{1.14}$$

where $E(\ )$ is the expenditure function, which minimizes the left-hand side of (1.7) as a function of $Q(K) \equiv 1/(1+r(K))$, q, and U. $E_2 \equiv \partial E/\partial q = G$ is the compensated demand function for G and $E_1 \equiv \partial E/\partial Q = c^2$ is the compensated demand function for c^2. It is assumed that each individual's expectation on the rate of interest for his old age is perfect foresight. This assumption is made only for the sake of simplicity. If we assume the static expectation, the analytical results would qualitatively be the same because we focus on the steady state properties of the system.

It is well known that in the voluntary provision of public good model redistribution of income does not matter at the interior solution (see, among others Warr, 1983). This neutrality result also holds in the present model. However, since our main concern is with intergenerational equity and the efficiency of the pension system, we do not investigate the issue of intragenerational redistribution by assuming that all interesting groups are identical.

We have assumed identical agents for analytical simplicity. If we allow for heterogeneity, the analytical framework would become complicated, but the main qualitative results below would hold as in the present model.

1.3 Cobb-Douglas example

In order to demonstrate concrete results with respect to dynamic properties, let us assume that the utility function (1.3) is given as a Cobb-Douglas one. The qualitative results are almost the same as in a more general production function.

$$U_t = (c_t^1)^{\alpha_1}(c_{t+1}^2)^{\alpha_2}(G_t)^{\alpha_3} \quad (\alpha_1 + \alpha_2 + \alpha_3 = 1) \tag{1.3$'$}$$

Then in this case we have

$$c_t^1 = \alpha_1 E_t \tag{1.15-1}$$

$$c_{t+1}^2 = \alpha_2(1 + r_{t+1})E_t \tag{1.15-2}$$

$$G_t = \alpha_3 E_t/q \tag{1.15-3}$$

From (1.2) (1.12) and (1.15), (1.14) may reduce to

$$nq\alpha G_t = K_{t+1} + \frac{1}{1 + r(K_{t+1})} n\beta w(K_{t+1}) \tag{1.16}$$

where $\alpha \equiv \alpha_2/\alpha_3$.

From (1.16) we have

$$G_t = G(K_{t+1}) \tag{1.17}$$

where

$$G'(K) = \frac{1}{nq\alpha}\{1 + A(K)\} > 0 \tag{1.18}$$

and

$$A(K) \equiv \frac{\beta\lambda(1-\lambda)n^\lambda K^{-\lambda}(1 + n^\lambda K^{-\lambda})}{(1 + (1-\lambda)n^\lambda K^{-\lambda})^2}.$$

From (1.18) we know $G'(\infty) = \frac{1}{nq\alpha}$ and $G'(0) = \frac{1+\beta\lambda}{nq\alpha}$. It is also easy to show $G'' < 0$.

G and K always move in the same direction, which implies a positive relation between G and K. Under the interdependence between the pension benefits and public spending, formulated as (1.4) and (1.5), an increase in public spending (or pension contribution) is consistent with economic growth. An increase in capital stock raises real income, stimulating the demand for public goods. In order to have a larger amount of public goods, it is necessary to reduce the subsidy from the general account to the pension account. Thus, the agent is willing to pay more pension contributions, making larger public spending possible. In general we have the substitution effect as well. Namely, an increase in K raises $1/(1+r)$, the intertemporal price of c^2, inducing a substitution from c^2 to G. This effect also produces larger pension contributions.

Hence, considering (1.10)(1.11)(1.12-1) and (1.12-3), (1.13) may be rewritten as

$$\frac{qn}{\alpha_3}G(K_{t+1}) = n(1-\beta)w(K_t) + \frac{n\beta}{[1+r(K_{t+1})]}w(K_{t+1}) + q(n-1)G(K_{t+1}) \quad (1.19)$$

Now we have

$$q(\theta n + 1 - n\alpha)G(K_{t+1}) + K_{t+1} = n(1-\beta)w(K_t) \quad (1.20)$$

where $\theta = (1-\alpha_3)/\alpha_3$. (1.20) may be expressed as

$$K_{t+1} = \Phi(K_t) \quad (1.20)'$$

which is the fundamental dynamic equation of the model.

Let us now investigate dynamics of (1.20) or (1.20)′. We have from (1.20)′

$$\Phi' = \frac{(1-\beta)q^{-1}n^{\lambda}K^{-\lambda}\lambda(1-\lambda)}{(\theta n + 1 - n\alpha)G' - q^{-1}} \quad (1.21)$$

The numerator when $K \Rightarrow \infty$

$$(\theta n + 1 - n\alpha)G' - \frac{1}{q} = \frac{1}{nq\alpha}(\theta n + 1)$$

is positive since $\theta > \alpha$ and hence $\theta n + 1 > n\alpha$. Thus, (1.21) is positive. We know that $\Phi'(0) > 1$.

The stability condition is

$$\Phi' = \frac{(1-\beta)q^{-1}n^{\lambda}K^{-\lambda}\lambda(1-\lambda)}{(\theta n + 1 - n\alpha)G' - q^{-1}} < 1$$

Or, we have

$$\Delta \equiv (\theta n - n\alpha + 1)G' - q^{-1} - (1-\beta)q^{-1}n^{\lambda}K^{-\lambda}\lambda(1-\lambda) > 0 \quad (1.22)$$

We assume

$$\theta n - 2\alpha n + 1 > 0$$

Then, we have that $\Phi'(\infty) < 1$. The system becomes dynamically stable. The larger the level of α_3 (the preference for public spending), it is more likely that the system would be stable.

1.4 Comparative statics

From (1.13) and (1.14) the steady-state equilibrium may be summarized by the following two equations.

$$nE[Q(K), q, U] = n[1-\beta+\beta Q(K)]w(K) + q(n-1)E_2[Q(K), q, U] \tag{1.23}$$

$$nE_1[Q(K), q, U] = \frac{1}{Q(K)}K + n\beta w(K) \tag{1.24}$$

Totally differentiating (1.23) and (1.24), we have

$$\frac{dU}{d\beta} = \frac{1}{\Omega}\left\{nw(-1+Q)\left[nE_{11}Q' - \beta w' - \frac{Q-Q'K}{Q^2}\right]\right.$$
$$\left. -[(nE_1 - q(n-1)E_{21})Q' - n(1-\beta+\beta Q)w' - n\beta wQ']nw\right\} \tag{1.25-1}$$

$$\frac{dK}{d\beta} = \frac{nw}{\Omega}[nE_U - q(n-1)G_U - (-1+\theta)nE_{1U}] < 0 \tag{1.25-2}$$

$$\frac{dU}{dq} = \frac{1}{\Omega}\left\{[bE_2 + q(n-1)E_{22}]\left[nE_{11}Q' - \beta w' - \frac{Q-Q'K}{Q^2}\right]\right.$$
$$\left. -(-nE_{12})[(nE_1 - q(n-1)E_{21})Q' - n(1-\beta+\beta Q)w' - n\beta wQ']\right\} \tag{1.25-3}$$

$$\frac{dK}{dq} = \frac{n}{\Omega}[-E_{12}[nE_U - q(n-1)G_U] - E_{1U}(n-1)(E_2 + qE_{22})] \tag{1.25-4}$$

where

$$\Omega \equiv [nE_U - q(n-1)G_U]\left[nE_{11}Q' - \beta w' - \frac{Q-KQ'}{Q^2}\right]$$
$$- nE_{1U}[(nE_1 - q(n-1)E_{21})Q' - n(1-\beta+\beta Q)w' - n\beta wQ']$$

We know $nE_U - q(n-1)E_{2U} > 0$, $E_{1U} > 0$. $nE_{11}Q' - \beta w' - \frac{Q-KQ'}{Q^2}$ is negative. And the sign of $(nE_1 - q(n-1)E_{21})Q' - n(1-\beta+\beta Q)w' - n\beta wQ'$ is ambiguous. When Ω becomes negative, an increase in exogenous resources will be beneficial, which is intuitively plausible. Thus, we assume $\Omega < 0$.

The effect of an increase in β or q on U is generally ambiguous. By setting (1.25-1) to be zero, we may implicitly derive the optimal level of β, the replacement ratio at the second best solution. An increase in β raises old-age consumption, which is desirable, while it reduces the effective income, the right-hand side of (1.23), which is undesirable. It is true that an increase in β raises old-age welfare, but it reduces young-age welfare. The overall welfare effect is ambiguous, which is intuitively plausible.

The sign of (1.25-3) is also ambiguous. It should thus be stressed that an increase in q could raise welfare at the second best solution. Intuition of this paradoxical result is as follows. On the one hand, an increase in q raises private savings due to the substitution effect, which may be beneficial. On the other hand, it reduces the (private) rate of return on contributions by raising the marginal cost of producing public goods, which income effect is not beneficial. The overall effect is thus ambiguous.

We know the sign of (1.25-2) becomes negative. As to the effect on capital accumulation, an increase in the replacement ratio will reduce accumulation of capital. The sign of (1.25-4) is ambiguous. On the one hand, an increase in q reduces the lifetime disposable income, producing the negative income effect on capital accumulation. On the other hand, an increase in q will reduce voluntary contributions, producing more savings due to the substitution effect. This stimulates capital accumulation. Thus, the overall effect on capital accumulation is ambiguous. In other words, an increase in the (private) rate of return of providing the contribution (a decreasing in q) may not necessarily stimulate accumulation of capital if the substitution effect is large.

The comparative static results suggest that there may exist a conflict between the first best and second best situations when q changes; a decrease in q is desirable at the first best solution but is not always desirable at the second best solution. The public sector does not have a strong incentive to raise the (private) rate of return of providing contributions when more capital accumulation is needed.

1.5 Normative aspects of public pension contributions

First best solution

In order to investigate the normative aspect of the model, it is useful to derive the first best solution. From (1.1)(1.2)(1.4)(1.5) and (1.9), the feasibility condition is given as

$$Y_t + K_t = K_{t+1} + n(c_t^1 + c_t^2) + qG_t \tag{1.26}$$

We analyze the optimal growth path which would be chosen by a central planner who maximizes an intertemporal social welfare function expressed as the sum of generational utilities discounted by the social discount factor on future generations, ρ, which is between 0 and 1. The discounted social objective is standard in the optimal growth literature, so I follow this. But it may be problematic because later generations are penalized.

$$Max \sum_{t=0}^{\infty} \rho^t U(c_t^1, c_{t+1}^2, G_t) \text{ subject to} \tag{1.26}$$

In other words, the first best problem is to maximize the Lagrange function

$$W = \sum_{t=0}^{\infty} \rho^t \{nU(c_t^1, c_{t+1}^2, G_t) - \mu_t[qG_t - Y_t - K_t + K_{t+1} + n(c_t^1 + c_t^2)]\} \tag{1.27}$$

where $\rho^t \mu_t$ is a Lagrange multiplier at time t.
The first-order conditions are as follows.

$$U_{1t} - \mu_t = 0 \tag{1.28-1}$$

$$U_{2t+1} - \mu_{t+1}\rho = 0 \tag{1.28-2}$$

$$nU_{3t} - q\mu_t = 0 \tag{1.28-3}$$

$$\mu_{t+1}(1 + r_{t+1})\rho - \mu_t = 0 \tag{1.28-4}$$

along with the transversality conditions

$$\lim_{t\to\infty} \rho^t \mu_t G_t = 0, \quad \lim_{t\to\infty} \rho^t \mu_t K_t = 0$$

where $U_{1t} = \partial U_t/\partial c_t^1$, $U_{2t+1} = \partial U_t/\partial c_{t+1}^2$, and $U_{3t} = \partial U_t/\partial G_t$.
From these conditions we have

$$\frac{U_{3t}}{U_{1t}} = \frac{q}{n} \tag{1.29-1}$$

$$\frac{U_{3t}}{U_{2t+1}} = \frac{q(r_{t+1} + 1)}{n} \tag{1.29-2}$$

From (1.29-1) and (1.29-2), we have

$$\frac{U_{3t}}{U_{1t}} + \frac{U_{3t}}{U_{2t+1}} < q + q(1 + r_{t+1}) \tag{1.30}$$

Note that $n > 1$. Since in the competitive economy we always have

$$\frac{U_{3t}}{U_{1t}} = q, \tag{1.31-1}$$

$$\frac{U_{3t}}{U_{2t+1}} = q(1 + r_{t+1}), \tag{1.31-2}$$

inequality (1.30) means that the public good-private consumption ratio $G_t/(c_t^1 + c_{t+1}^2)$ in the competitive economy is smaller than in the first best economy. Each agent does not fully recognize the total effect of voluntary contributions to public spending. That is, the main reason for such under-provision of G in the relative sense is that each group disregards a positive externality of cooperation with public pension contribution which spills over into all other groups in choosing its own contribution [This is the conventional result in the literature. See Bliss and Nalebuff (1984), Bergstrom et al. (1986), Boadway et al. (1989) and Ihori and Itaya (2002)(2003).] The pension contribution (or public spending) divided by private consumption is too little in the pension system with public subsidies at the second best solution. In this sense, the pension contribution is too little and private consumption is too much in the competitive economy in terms of the static efficiency.

In terms of the dynamic efficiency from (1.28-4) we have as the modified golden rule:

$$(1 + r)\rho = 1 \tag{1.32}$$

The first best solution may be summarized by (1.26), (1.29-1), (1.29-2) and (1.32).

Under the Cobb-Douglas utility function (1.3)′, (1.29-1) and (1.29-2) are rewritten as

$$\frac{\alpha_3 c^1}{\alpha_1 G} = \frac{q(r+1)}{(1+r)n} \tag{1.29-1$'$}$$

$$\frac{\alpha_3 c^2}{\alpha_2 G} = \frac{q(r+1)}{n} \tag{1.29-2$'$}$$

Substituting these two equations into (1.26) and considering (1.32), we obtain the first best G, G_{FB}, as

$$G_{FB} = \frac{Y_{FB}}{\dfrac{q}{\alpha_3}\dfrac{1}{\rho}(\rho\alpha_1 + \alpha_2) + q} \tag{1.33}$$

where Y_{FB} is output associated with the modified golden rule. The first best pension contribution is decreasing with q and increasing with ρ and α_3. These results are intuitively plausible. For example, the larger the concern for future generations (ρ), the larger the first best levels of capital accumulation and public goods.

Steady-state level of pension contribution

In the standard overlapping generations growth model it is well known that capital may be over-supplied in the competitive equilibrium. Capital may be too much in this model as well when the competitive steady-state economy is on the inefficient path ($(1+r)\rho < 1$), and vice versa. The smaller the concern for the future generations, it is more likely to have the inefficient case.

Let us compare the steady-state levels of pension contribution or public spending, G^* at the Nash solution given by (1.23)(1.24), and G_{FB} at the first best solution given by (1.33). Remember that the pension contribution at the second best solution is independent of ρ, while G_{FB} is increasing with ρ. Hence, the steady-state level of pension contribution may be too much if q is very high or ρ is very small. If $r^* \geq \frac{1-\rho}{\rho}$, it is easy to see $G^* < G_{FB}$. However, if $r^* < \frac{1-\rho}{\rho}$, we cannot exclude the possibility of $G^* > G_{FB}$. Pension contributions at the non-cooperative Nash solution could be over-provided. In other words, G^* given by (1.23)(1.24) could be higher than G_{FB} given by (1.33) if ρ is small enough. If α_2 and q are large, we could have the same possibility. The pension contribution at the second best solution may be increasing with q, while its steady-state level at the first best solution is decreasing with q. Thus, when q is high, it is likely that the pension contribution is too much even if each interest group can easily avoid contributions.

In Figure 1.1 line AB represents the feasibility condition at the steady state where the modified golden rule (1.32) is satisfied.

$$n(c^1 + c^2) - bG = Y_{FB} \tag{1.34}$$

Point F is the first best solution, while point N is the Nash equilibrium when $r^* = \frac{1-\rho}{\rho}$. The movement from F to N on the same budget line AB reflects the free-riding effect. As shown in Figure 1.1, G is too little at point N compared with point F.

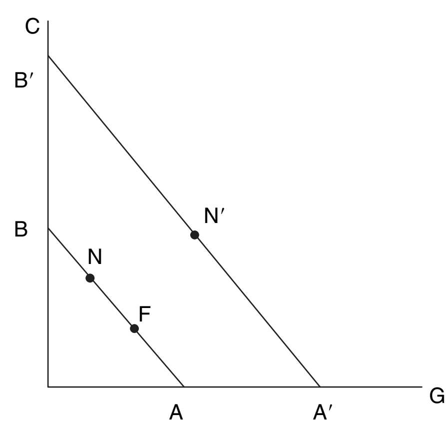

Figure 1.1 Steady-state level of pension contributions

If $r^* < \frac{1-\rho}{\rho}$, we may draw Figure 1.1 where line A′B′ represents the feasibility condition at the steady state in the competitive economy. A shift from line AB to line A′B′ reflects the income effect. Thus, if the income effect dominates the free-riding effect, G^* at point N′ could be greater than G_{FB} at point F. The lower the discount factor ρ and the concern with future consumption α_2, it is more likely to have such a paradoxical case. In such a case the pension contribution is too much even if each interest group can easily avoid contributions and the government cannot effectively penalize the avoiding behaviour.

If we allow for positive population growth, then the modified golden rule (1.32) will be altered to

$$(1+r)\rho = 1+\gamma \tag{1.32$'$}$$

where γ is the population growth rate. In this case the competitive steady-state economy is on the efficient path if and only if $(1+r)\rho > 1+\gamma$. Hence, even if ρ is close to 1, it is still possible that the path is inefficient where $\gamma > r$, and we cannot exclude the possibility of $G^* > G_{FB}$.

Optimal tax and subsidy policy

Finally, let us consider some tax and subsidy policy to attain the first best economy in the long run. Suppose that consumption taxes η^1, η^2, a subsidy to contributions ε and a tax on interest income μ are available. Then, the consumer's budget constraints (1.1) and (1.2) are rewritten as

$$(1+\eta^1)c_{it}^1 = (1-\tau)w_{it} - (1-\varepsilon)g_{it} - s_{it} \tag{1.1$'$}$$

$$(1+\eta^2)c_{it+1}^2 = [1+(1-\mu)r_{t+1}]s_{it} + \beta w_{t+1} \tag{1.2$'$}$$

Then, considering (1.1)′ and (1.2)′, the first-order conditions (1.31-1) and (1.31-2) are rewritten as

$$\frac{U_{3t}}{U_{1t}} = \frac{q(1-\varepsilon)}{1+\eta^1}, \tag{1.3$'$}$$

$$\frac{U_{3t}}{U_{2t+1}} = \frac{q[1+(1-\mu)r_{t+1}]}{1+\eta^2}, \tag{1.4$'$}$$

From (1.29-1)(1.29-2) and (1.31-1)′(1.31-2)′, the optimal conditions are given as

$$\frac{q(1-\varepsilon)}{1+\eta^1} = \frac{q}{n}$$

$$\frac{q[1+(1-\mu)r]}{1+\eta^2} = \frac{q(1+r)}{n}$$

Hence, the optimal values of $\varepsilon, \mu, \eta^1, \eta^2$ are not uniquely determined. For example, when $\eta^1 = \eta^2 = 0$, the optimal values of ε and μ are respectively given as

$$\varepsilon^* = 1 - \frac{1}{n} > 0 \tag{1.35-1}$$

$$\mu^* = \frac{(1+r)(n-1)}{rn} \tag{1.35-2}$$

(1.35-1) means that a subsidy to contributions is used for attaining the static efficiency. In Japan public pension contributions are fully exempted from the income tax base. It actually subsidizes contributions. Such a subsidy to contributions may be justified to realize the static efficiency although the actual level of subsidies may not be optimal. From (1.35-2) it is desirable to tax interest income to stimulate pension contributions. Note that a tax on old-age benefits is not effective since it cannot affect the first order conditions of consumers.

Alternatively, when $\varepsilon = \mu = 0$, the optimal values of consumption taxes are respectively given as,

$$\eta^{1*} = \eta^{2*} = n - 1 > 0 \tag{1.36}$$

The optimal consumption tax rates are uniform and positive. Intuition is as follows. By taxing private consumption, providing pension contributions so as to supply public goods becomes more favourable. It would thus stimulate pension contributions, which is desirable to internalize the free-riding behaviour of interest groups.

As to the dynamic efficiency, intergenerational redistribution policy would be useful. Or if the government can control the replacement ratio, it would attain the dynamic efficiency by realizing the modified golden rule. From (1.25-2) the replacement ratio affects capital accumulation negatively. So the government may choose the replacement ratio to attain (1.32). When physical capital is over-provided at the Nash equilibrium ($r^* \leq \frac{1-\rho}{\rho}$), the intergenerational transfer from the young to the old would be desirable. An increase in the replacement ratio could have this effect. Such a policy would reduce the lifetime disposable income, reducing savings.

On the contrary, when capital is under-provided ($r^* > \frac{1-\rho}{\rho}$), the intergenerational transfer from the old to the young such as a decrease in the replacement ratio would be desirable. In this case the pension contribution is too little due to the under-accumulation of capital, and it is thus necessary to stimulate savings. The larger the concern for future generations, it is more likely to have this case.

1.6 Conclusion

Suppose the pay-as-you-go system has to be maintained as the means of intergenerational transfer. Without effective enforcement measures to collect pension contributions from various interest groups, the government would face the difficulty of maintaining the pay-as-you-go system in an ageing society. This chapter has shown that interest groups 'voluntarily' provide a social security contribution if a part of pension benefits is financed by subsidies from the general account of national government. This is so because, under this subsidy mechanism each interest group has an incentive to contribute to the pay-as-you-go pension system even if the population growth rate is less than the real rate of interest.

An increase in capital stock raises real income, stimulating the demand for public goods. In order to have a larger amount of public goods, it is necessary to raise pension contributions. Thus, the agent is willing to accept more pension contributions, resulting in larger public spending. The larger pension contribution and capital accumulation would be likely to coexist at the competitive solution. In this sense, the subsidy mechanism is a realistic compromise to let each interest group cooperate with the otherwise unpopular pay-as-you-go system in the situation of an ageing population.

We have also clarified how the relevant parameters would affect dynamic properties. An increase in the replacement ratio will reduce accumulation of capital, although its effect on welfare is ambiguous. There may exist a conflict between the first best and decentralized situations when the cost of public goods changes. The public sector does not have a strong incentive to raise the (private) productivity of providing pension contribution although it benefits all generations at the first best solution.

It is well known that capital may be too much in the long run when the competitive steady-state economy is on the inefficient path. Even if the concern for future generations are large, it is still possible to have such an inefficient case. We have shown that in such a case, the steady-state level of pension contributions may be too much in the long run, although it is too little compared with consumption. The pension contribution could be too much even if each interest group can easily avoid contributions and the government cannot penalize the avoiding behaviour.

As to attaining the static efficiency, consumption taxes or combination of a subsidy to social security contributions and an interest income tax would be useful for correcting the free-riding behaviour of interest groups. This is desirable even if capital is over-accumulated. If the government can control the replacement ratio, it can attain the dynamic efficiency by realizing the modified golden rule. When capital is under-provided, the intergenerational transfer from the old to the young such as a decrease in the replacement ratio would be desirable. The larger the concern for future generations, the more likely it is to have this case.

Notes

1. Abel (1999) develops a tractable stochastic overlapping generations model to analyze the equilibrium equity premium and growth rate of the capital stock in the presence of a defined-benefit social security system.
2. From the 2004 pension reform Japan is moving from the defined benefit system to the defined contribution system. See Takayama (2007).
3. As for voluntary provision of public goods, see Shibata (1971), Warr (1983), and Bergstrom et al. (1986) among others.

References

Abel, A.B. (1999) 'The Social Security Trust Fund, the Riskless Interest Rate and Capital Accumulation', NBER Working Paper no. 6991.

Auerbach, A. and L. Kotlikoff (1987) *Dynamic Fiscal Policy* (Cambridge: Cambridge University Press).

Bergstrom, T., L. Blume and H. Varian (1986) 'On the Private Provision of Public Goods', *Journal of Public Economics*, vol. 29, pp. 25–49.

Bliss, C. and B. Nalebuff (1984) 'Dragon-slaying and Ballroom Dancing: the Private Supply of a Public Good', *Journal of Public Economics*, vol. 25, pp. 1–12.

Boadway, R., P. Pestieau and D. Wildasin (1989) 'Tax-transfer Policy and the Voluntary Provision of Public Goods', *Journal of Public Economics*, vol. 39, pp. 157–76.

Cifuentes, R. and S. Valdes-Prieto (1997) 'Transitions in the Presence of Credit Constraints', in S. Valdes-Prieto (ed.), *The Economics of Pensions* (Cambridge: Cambridge University Press).

Hatta, T. and T. Oguchi (1999) *Nenkinkaikakuron (Japanese Pension Reform)* (Tokyo: Nihonkeizai-shinbunsya). In Japanese.

Ihori, T. (2002) 'Pension Contributions and Capital Accumulation under a Modified Funded System', in T. Ihori and T. Tachibanaki (eds), *Social Security Reform in Advanced Countries* (London: Routledge).

Ihori, T. and J. Itaya (2002) 'A Dynamic Model of Fiscal Reconstruction', *European Journal of Political Economy*, vol. 17, pp. 1057–97.

Ihori, T. and J. Itaya (2003) 'Fiscal Reconstruction and Local Government Financing', *International Tax and Public Finance*, vol. 10, pp. 1–13.

Mulligan, C.B. and X. Sala-i-Martin (1999a) 'Social Security in Theory and Practice (I): Facts and Political Theories', NBER Working Paper no. 7118.

Mulligan, C.B. and X. Sala-i-Martin (1999b) 'Social Security in Theory and Practice (2): Efficiency Theories, Narrative Theories and Implications for Reform', NBER Working Paper no. 7119.

Oshio, T. (2004) 'Social Security and Trust Fund Management', NBER Working Paper no. 10444.

Shibata, H. (1971) 'A Bargaining Model of the Pure Theory of Public Expenditure', *Journal of Political Economy*, vol. 79, pp. 1–29.

Takayama, N. (1998) *The Morning After in Japan: Its Declining Population, Too Generous Pensions and a Weakened Economy* (Tokyo: Maruzen).

Takayama, N. (2007) 'Social Security Pensions and Intergenerational Equity: the Japanese Case', chapter 3 in this volume.

Warr, P.G. (1983) 'The Private Provision of a Public Good Is Independent of the Distribution of Income', *Economics Letters*, vol. 13, pp. 207–11.

2

Equity and Efficiency in Overlapping Generations Economies

Tomoichi Shinotsuka, Koichi Suga, Kotaro Suzumura and Koichi Tadenuma

2.1 Introduction

This chapter studies equity and efficiency of allocations in the overlapping generations economy formulated by Samuelson (1958). A central notion of distributional equity is *no-envy* (Foley, 1967; Kolm, 1972): no person prefers the consumption of any other person to his/her own. Suzumura (2002) proposed three distinct notions of equity based on no-envy in overlapping generations economies. The first one concerns contemporary (overlapping) consumptions in each time period. It requires that for each period, no person should prefer the consumption of any other person in that period to his/her own. In contrast, the second notion is about lifetime consumption plans. It stipulates that no person should prefer the lifetime consumption plan of any other person to his/her own. The third notion is based on the *lifetime rate of return* due to Cass and Yaari (1966). It simply requires an equal lifetime rate of return for all persons.

In a simple model where there is one (composite) commodity with no production, the preferences are identical, and the rate of population growth is constant, we examine the implications of each notion of equity, and clarify the logical relations of the three notions, paying special attention to their relations with stationarity of allocations. We also study the existence and characterizations of allocations attaining maximal utility under no-envy in lifetime consumption plans, which are the allocations selected by the *equity-first and efficiency-second principle* due to Tadenuma (2002).

The rest of the chapter is organized as follows. The next section formulates the model and the notions of equity. The following three sections examine the implications of no-envy in contemporary (overlapping) consumptions, equality in lifetime rate of return, and no-envy in lifetime consumption plans, respectively. Section 6 studies allocations attaining maximal utility under no-envy in lifetime consumption plans. The final section summarizes the results.

2.2 The model and the notions of equity

We consider the overlapping generations model formulated by Samuelson (1958). Each person lives for two periods, earning one unit of a perishable good in the first period, and earning nothing in the second period. Let C^i denote an amount of consumption in his i-th period. Preferences are assumed to be identical for all persons. Let $U : \mathbb{R}^2_+ \to \mathbb{R}$ be the common utility function, which is increasing in each argument. The value $U(C^1, C^2)$ represents the utility level when a person consumes C^1 and C^2 in his first period and in his second period, respectively. At the beginning of each period t, $(1+n)^t$ persons are born where $n \geq 0$ is the *rate of population growth*. We call the persons born at t *generation t*.

In order to focus on *inter*-generational distribution problems rather than *intra*-generational problems, we assume that all persons belonging to the same generation consumes the same bundle. Let $\mathbf{C}_t = (C^1_t, C^2_t)$ be a lifetime consumption plan of each person in generation t. Since $(1 + n)^t$ persons are born at the beginning of each period t, the aggregate consumption vector of generation t is given by $(1 + n)^t\mathbf{C}_t$. An *allocation* is a doubly infinite and nonnegative sequence $\{\mathbf{C}_t\}_{t\in Z}$, where Z denotes the set of all integers. An allocation $\{\mathbf{C}_t\}_{t\in Z}$ is *stationary* if $\mathbf{C}_{t-1} = \mathbf{C}_t$ for all $t \in Z$. An allocation $\{\mathbf{C}_t\}_{t\in Z}$ is *feasible* if for every $t \in Z$,

$$C^1_t + \frac{C^2_{t-1}}{1+n} \leq 1, \tag{2.1}$$

and it is *exactly feasible* if for every $t \in Z$,

$$C^1_t + \frac{C^2_{t-1}}{1+n} = 1. \tag{2.2}$$

An allocation $\{\mathbf{C}_t\}_{t\in Z}$ is *Pareto efficient* if it is feasible and there exists no feasible allocation $\{\mathbf{C}'_t\}_{t\in Z}$ such that $U(\mathbf{C}'_t) \geq U(\mathbf{C}_t)$ for all $t \in Z$ with strict inequality for some t.

Following Suzumura (2002), we formulate three distinct notions of equity-as-no-envy. The first one requires that for each time period, no person prefers the consumption of any other person in the same period. Since there is only one (composite) commodity in our model, the condition reduces to the following one:

No-Envy in Overlapping Consumptions (NEOC): For all $t \in Z$, $C^2_{t-1} = C^1_t$.

The second notion means that no person prefers the lifetime consumption plan of any other person to his/her own. Because it is assumed that the preferences are identical, and that all persons belonging to the same generation have the same lifetime consumption, the condition reduces to the following one:

No-Envy in Lifetime Consumptions (NELC): For all $t \in Z$, $U(\mathbf{C}_{t-1}) = U(\mathbf{C}_t)$.

The third notion of equity is based on the *lifetime rate of return* due to Cass and Yaari (1966). Consider allocations $\{\mathbf{C}_t\}_{t \in Z}$ such that $C_t^1 > 0$ and $C_t^2 > 0$ for all t. We call such allocation *positive*. Note that exact feasibility and positivity together imply $C_t^1 < 1$ for all $t \in Z$. Thus, the following notion of lifetime rate of return is well-defined:

$$r_t = \frac{C_t^2 - (1 - C_t^1)}{1 - C_t^1}.$$

Equity in Lifetime Rate of Return (ELRR): For all positive and exactly feasible allocations, and for all $t \in Z$, $r_{t-1} = r_t$.

In the following sections, we investigate implications and logical relations of these equity notions.

2.3 No-envy in overlapping consumptions

We start with examining implications of No-Envy in Overlapping Consumptions (NEOC). It turns out that this is the strongest requirement of all the equity notions introduced above, and a fundamental trade-off between equity and efficiency emerges.

Proposition 1 *There exists one and only one allocation satisfying exact feasibility and NEOC. It is given by*

$$\hat{\mathbf{C}}_t = \hat{\mathbf{C}} := \left(\frac{n+1}{n+2}, \frac{n+1}{n+2}\right)$$

for all $t \in Z$. If the marginal rate of substitution of consumption in period 2 for consumption in period 1 at $\hat{\mathbf{C}}$ is not equal to $1+n$, then the allocation is not Pareto efficient.

Proof: Let $\{\hat{\mathbf{C}}_t\}_{t \in Z}$ be an allocation satisfying exact feasibility and NEOC. Let $t \in Z$ be given. By exact feasibility,

$$\hat{C}_t^1 + \frac{\hat{C}_{t-1}^2}{1+n} = 1. \tag{2.3}$$

By NEOC,

$$\hat{C}_{t-1}^2 = \hat{C}_t^1 \tag{2.4}$$

Solving equations (2.3) and (2.4), we have

$$(\hat{C}_t^1, \hat{C}_{t-1}^2) = \left(\frac{n+1}{n+2}, \frac{n+1}{n+2}\right).$$

Because the rate of population growth n is constant, the solution does not depend on t. Hence, $\hat{C}^2_{t-1} = \hat{C}^2_t$. Therefore, for all $t \in Z$,

$$\hat{\mathbf{C}}_t := (\hat{C}^1_t, \hat{C}^2_t) = \left(\frac{n+1}{n+2}, \frac{n+1}{n+2}\right) = \hat{\mathbf{C}}.$$

This means that the allocation $\{\hat{\mathbf{C}}_t\}_{t\in Z}$ is stationary.

Define

$$\eta(\hat{\mathbf{C}}) = \frac{\frac{\partial}{\partial C^1} U(\hat{\mathbf{C}})}{\frac{\partial}{\partial C^2} U(\hat{\mathbf{C}})}.$$

$\eta(\hat{\mathbf{C}})$ is the marginal rate of substitution of consumption in period 2 for consumption in period 1 at $\hat{\mathbf{C}}$. Notice that a stationary allocation $\{\mathbf{C}_t\}_{t\in Z}$ is feasible if and only if

$$C^1_t + \frac{C^2_t}{1+n} \leq 1$$

for all $t \in Z$. Hence, if $\eta(\hat{\mathbf{C}}) \neq 1+n$, then there exists a feasible and stationary allocation that Pareto dominates $\{\hat{\mathbf{C}}_t\}_{t\in Z}$ (Figure 2.1). *Q.E.D.*

Let us call the allocation $\{\hat{\mathbf{C}}_t\}_{t\in Z}$ defined above the *NEOC allocation*.

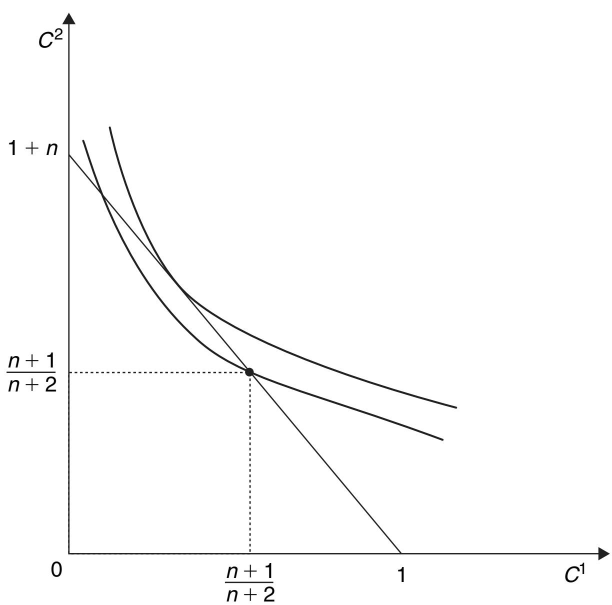

Figure 2.1 Trade-off between NEOC and Pareto efficiency

Corollary 1 *If an exactly feasible allocation satisfies NEOC, then it also satisfies both NELC and ELRR.*

Proof: From Proposition 1, if an exactly feasible allocation satisfies NEOC, then it must be the NEOC allocation. Since the NEOC allocation is stationary, it satisfies both NELC and ELRR. *Q.E.D.*

Corollary 2 *If the marginal rate of substitution of consumption in period 2 for consumption in period 1 at* $\hat{\mathbf{C}}$ *is not equal to* $1+n$*, then there exists no allocation satisfying NEOC and Pareto efficiency together.*

Proof: Suppose, on the contrary, that there exists an allocation satisfying NEOC and Pareto efficiency together. Since it is Pareto efficient, it must be exactly feasible. By Proposition 1, exact feasibility and NEOC imply that the allocation is the NEOC allocation. But since $\eta(\hat{\mathbf{C}}) \neq 1+n$, it also follows from Proposition 1 that the allocation cannot be Pareto efficient. This is a contradiction. *Q.E.D.*

Consider the separable utility function U defined by

$$U(C^1, C^2) = u(C^1) + \delta u(C^2) \tag{2.5}$$

where δ is a positive constant and u is a twice continuously differentiable function with a positive first derivative and a negative second derivative. Then, $\delta^{-1} - 1$ is called the *pure rate of time preference*. In this case,

$$\eta(\hat{\mathbf{C}}) = \frac{\frac{\partial}{\partial C^1} U(\hat{\mathbf{C}})}{\frac{\partial}{\partial C^2} U(\hat{\mathbf{C}})} = \delta^{-1}.$$

Hence, we have the following corollary.

Corollary 3 *In the case of the separable utility function, if the pure rate of time preference is not equal to the rate of population growth, then there exists no allocation satisfying NEOC and Pareto efficiency together.*

Proposition 1 has two interesting implications. First, *growth is compatible with equity*. If there were no growth, each young person would have to split her earnings to give a half to an old person in order to achieve equity in overlapping consumptions. The faster the economy grows, the more each person can receive because she has to give less to older persons in her young age while she can receive more from younger persons in her old age. Indeed, the NEOC allocation is strictly increasing in the rate of population growth n.

The second implication of Proposition 1 is that *there exists almost surely a room for improvement in the welfare of all persons at the NEOC allocation*. In the static model with one commodity, equal split is always Pareto efficient as well as equitable. By contrast, in this simple model with overlapping generations, equal split among contemporary persons at each period does not lead to a Pareto efficient allocation.

2.4 Equity in lifetime rate of return and the biological rate of interest

In this section, we consider positive and exactly feasible allocations, for which the lifetime rate of return r_t is well-defined.

It is easy to see that the lifetime rate of return associated with the NEOC allocation is equal to n. That is, *NEOC implies that the lifetime rate of return should be equal to the 'biological rate of interest'*. This welfare implication on the biological rate of interest seems new.

As we saw in the previous section, under exact feasibility, NEOC implies Equity in Lifetime Rate of Return (ELRR), but the converse is not true. Indeed, any stationary allocation satisfies ELRR. Hence, interesting questions would be:

(1) Are there non-stationary allocations satisfying ELRR?
(2) Does ELRR provide some implications on the biological rate of interest?

Our next result answers the above questions.

Proposition 2 *Let $\{\mathbf{C}_t\}_{t\in Z}$ be an exactly feasible, positive allocation satisfying ELRR with the common lifetime rate of return r being nonnegative. Then, $r = n$ and the allocation is stationary.*

Proof: By ELRR, $C_t^2 = (1+r)(1-C_t^1)$ for all t. By exact feasibility, $C_{t-1}^2 = (1+n)(1-C_t^1)$ for all t. Hence, $(1+r)(1-C_t^1) = (1+n)(1-C_{t+1}^1)$ for all t. Let $\lambda = (1+r)/(1+n)$. Since r is nonnegative, λ is positive. Then, $(1-C_{t+1}^1) = \lambda(1-C_t^1)$ for all t. Hence, $(1-C_t^1) = \lambda^t(1-C_0^1)$ for all t. Note that $1-C_t^1$ is positive for all t since $\{\mathbf{C}_t\}_{t\in Z}$ is exactly feasible and positive. If $\lambda > 1$, $\{|1-C_t^1|\}_{t\in Z}$ goes to infinity as t goes to infinity. If $\lambda < 1$, $\{|1-C_t^1|\}_{t\in Z}$ goes to infinity as t goes to *minus infinity*. Hence, exact feasibility is violated unless $\lambda = 1$. Therefore, $r = n$. Then, the allocation is stationary. *Q.E.D.*

Proposition 2 means that, for positive and exactly feasible allocations, *ELRR is equivalent to stationarity*. In other words, an allocation satisfies ELRR if and only if the lifetime rate of return of each generation is equal to the biological rate of interest. This seems to be an interesting characterization of the biological rate of interest.

In Figure 2.1, the set of all exactly feasible and positive allocations satisfying ELRR are depicted as the line

$$C^1 + \frac{C^2}{1+n} = 1.$$

The point $((n+1)/(n+2), (n+1)/(n+2))$ on the line represents the unique (under exact feasibility) allocation satisfying NEOC. This figure clearly shows that under exact feasibility, NEOC implies ELRR.

2.5 No-envy in lifetime consumptions

In this section, we investigate implications of requiring No-Envy in Lifetime Consumptions (NELC). Clearly, any stationary allocation satisfies NELC. The question is: *Is there any non-stationary allocation satisfying exact feasibility and NELC?*

A basic observation is that exact feasibility and NELC together generate a dynamical system defined by the following difference equation: for all $t \in Z$,

$$U(C_{t-1}^1, (1+n)(1-C_t^1)) = U(C_t^1, (1+n)(1-C_{t+1}^1)). \tag{2.6}$$

By nonnegativity of consumption and exact feasibility,

$$0 \le C_t^1 \le 1 \tag{2.7}$$

for all $t \in Z$.

Proposition 3 *Let $\{\mathbf{C}_t\}_{t \in Z}$ be an exactly feasible allocation satisfying NELC. If it is not a stationary allocation, then there exist $\bar{C}^1 \in [0,1]$ and $\bar{\bar{C}}^1 \in [0,1]$ such that* $\lim_{t\to\infty} C_t^1 = \bar{C}^1$, $\lim_{t\to-\infty} C_t^1 = \bar{\bar{C}}^1$, *and for all $t \in Z$,* $U(\mathbf{C}_t) = U(\bar{C}^1, (1+n)(1-\bar{C}^1)) = U(\bar{\bar{C}}^1, (1+n)(1-\bar{\bar{C}}^1))$.

Proof: Assume that $\{\mathbf{C}_t\}_{t \in Z}$ is exactly feasible, satisfies NELC, but is not stationary. Then, there exists $t^* \in Z$ such that $C_{t^*-1}^1 \ne C_{t^*}^1$.

Case 1: $C_{t^*-1}^1 < C_{t^*}^1$.

By equation (2.6) and the strict monotonicity of U, it must be true that $(1+n)(1-C_{t^*}^1) > (1+n)(1-C_{t^*+1}^1)$. Hence, $C_{t^*}^1 < C_{t^*+1}^1$. Repeating this argument, we have $C_{t^*-1}^1 < C_{t^*}^1 < C_{t^*+1}^1 < C_{t^*+2}^1 \ldots$. That is, C_t^1 is monotonically increasing as t increases. Since C_t^1 is bounded in $[0,1]$, there exists $\bar{C}^1 \in [0,1]$ such that $\lim_{t\to\infty} C_t^1 = \bar{C}^1$.

On the other hand, by $C_{t^*-1}^1 < C_{t^*}^1$, we have $(1+n)(1-C_{t^*-1}^1) > (1+n)(1-C_{t^*}^1)$. It follows from equation (2.6) and the strict monotonicity of U that $C_{t^*-2}^1 < C_{t^*-1}^1$. Repeating this argument, we can show that C_t^1 is monotonically decreasing as t decreases. Since C_t^1 is bounded in $[0,1]$, there exists $\bar{\bar{C}}^1 \in [0,1]$ such that $\lim_{t\to-\infty} C_t^1 = \bar{\bar{C}}^1$.

By continuity of U and equation (2.6), we have for all $t \in Z$, $U(\mathbf{C}_t) = U(\bar{C}^1, (1+n)(1-\bar{C}^1)) = U(\bar{\bar{C}}^1, (1+n)(1-\bar{\bar{C}}^1))$.

Case 2: $C^1_{t^*-1} > C^1_{t^*}$.

It can be shown that C^1_t is monotonically decreasing as t increases, and monotonically increasing as t decreases. Then, the claim in the proposition follows from a similar argument to case 1. *Q.E.D.*

Linear utility case

In this subsection, we assume that the utility function U is linear:

$$U(C^1, C^2) = aC^1 + C^2 \tag{2.8}$$

where a is a positive constant.

Proposition 4 *If the utility function U is linear, then there does not exist a non-stationary allocation satisfying exact feasibility and NELC.*

Proof: Substituting (2.8) into (2.6), we obtain the following linear, second order difference equation:

$$aC^1_{t-1} + (1+n)(1 - C^1_t) = aC^1_t + (1+n)(1 - C^1_{t+1}). \tag{2.9}$$

Rearranging the terms gives

$$(1+n)C^1_{t+1} - (1+n+a)C^1_t + aC^1_{t-1} = 0. \tag{2.10}$$

Consider the corresponding characteristic equation:

$$(1+n)x^2 - (1+n+a)x + a = 0. \tag{2.11}$$

The solutions to (2.11) are:

$$x_1 = 1,\ x_2 = \frac{a}{1+n}. \tag{2.12}$$

Therefore, the solution to the difference equation (2.6) is as follows:
if $a = 1+n$,

$$C^1_t = B_1 + B_2 t, \tag{2.13}$$

and if $a \neq 1+n$,

$$C^1_t = B_1 + B_2 x_2^t, \tag{2.14}$$

where B_1 and B_2 are constants depending on the initial condition. Note that an allocation $\{\mathbf{C}_t\}_{t\in Z}$ is stationary if and only if $\{C^1_t\}_{t\in Z}$ is constant over time.

The case of $a = 1+n$ is divided into three subcases. If $B_2 > 0$, then $\{C^1_t\}_{t\in Z}$ goes above 1 as t goes to infinity, and it goes below 0 as t goes to minus infinity. Thus the feasibility condition does not hold for sufficiently large or small t. If $B_2 < 0$, then $\{C_t^1\}_{t \in Z}$ goes below 0 as t goes to infinity, and it goes above 1 as t goes to minus infinity. Thus the feasibility condition does not hold for sufficiently large or small t. When $B_2 = 0$, $\{C_t^1\}_{t \in Z}$ is constant.

Suppose $\frac{a}{1+n} < 1$. If $\{C_t^1\}_{t \in Z}$ is not constant over time, then $\{C_t^1\}_{t \in Z}$ goes to infinity as t goes to minus infinity. Hence, there does not exist a non-stationary allocation satisfying exact feasibility and NELC.

Suppose $\frac{a}{1+n} > 1$. If $\{C_t^1\}_{t \in Z}$ is not constant over time, then $\{C_t^1\}_{t \in Z}$ goes to infinity as t goes to infinity. Hence, there does not exist a non-stationary allocation satisfying exact feasibility and NELC in this case either. *Q.E.D.*

Quasi-linear utility case

In this subsection, we assume that the utility function U is of the following form:

$$U(C^1, C^2) = v(C^1) + C^2, \tag{2.15}$$

where $v(C^1)$ is twice-continuously differentiable and has positive first derivative and negative second derivative. Define $\alpha := \min \{v'(c) \mid c \in [0, 1]\}$ and $\beta := \max \{v'(c) \mid c \in [0, 1]\}$. The parameters α and β measure the magnitudes of marginal utility in the first-period consumption.

Proposition 5 *If either $\alpha > 1 + n$ or $\beta < 1 + n$, there does not exist a non-stationary allocation satisfying exact feasibility and NELC.*

Proof: Substituting (2.15) into (2.6), we obtain a non-linear, second order difference equation:

$$v(C_{t-1}^1) + (1 + n)(1 - C_t^1) = v(C_t^1) + (1 + n)(1 - C_{t+1}^1). \tag{2.16}$$

Rearranging the terms, we have

$$C_{t+1}^1 - C_t^1 = \frac{1}{1+n}[v(C_t^1) - v(C_{t-1}^1)]. \tag{2.17}$$

By the mean value theorem, there exists $c_t^* \in [0, 1]$ such that

$$|C_{t+1}^1 - C_t^1| = \frac{v'(c_t^*)}{1+n}|C_t^1 - C_{t-1}^1|. \tag{2.18}$$

Assume that $\alpha > 1 + n$. For all $t \in Z$, since $v'(c_t^*) \geq \alpha > 1 + n$, we have $\frac{v'(c_t^*)}{1+n} > 1$. It then follows from (2.18) that if $\{C_t^1\}_{t \in Z}$ is not constant, then $\{|C_{t+1}^1 - C_t^1|\}_{t \in Z}$ goes to infinity as t goes to infinity. Therefore, $\{C_t^1\}_{t \in Z}$ violates feasibility.

Next, assume that $0 < \beta < 1 + n$. Then, $0 < \frac{\beta}{1+n} < 1$. For all $t \in Z$, since $v'(c_t^*) \leq \beta$, it follows from (2.18) that

$$|C_{t+1}^1 - C_t^1| \leq \frac{\beta}{1+n}|C_t^1 - C_{t-1}^1|.$$

By iteration, for all positive integers T,

$$|C_{t+1}^1 - C_t^1| \leq \left(\frac{\beta}{1+n}\right)^T |C_{t-T+1}^1 - C_{t-T}^1|.$$

Letting $t = 0$, for all positive integers T, we obtain

$$\left(\frac{\beta}{1+n}\right)^{-T} |C_1^1 - C_0^1| \leq |C_{-T+1}^1 - C_{-T}^1|.$$

Therefore, if $\{C_t^1\}_{t\in Z}$ is not constant, then $\{|C_{t+1}^1 - C_t^1|\}_{t\in Z}$ goes to infinity as t goes to *minus infinity*. *Q.E.D.*

Non-stationary allocation satisfying NELC

If the utility function is not linear or quasi-linear, then NELC does not imply stationarity. There exist non-stationary allocations satisfying NELC, as we present an example in what follows.

Let $n > -1$ be the rate of population growth. Let $a > \max\{1 + n, \frac{1}{1+n}\}$ be given. The utility function U is defined by:

$$U(C^1, C^2) = \begin{cases} aC^1 + C^2 & \text{if } C^1 \leq C^2 \\ C^1 + aC^2 & \text{if } C^1 > C^2. \end{cases}$$

Since $a > \max\left\{1 + n, \frac{1}{1+n}\right\}$, we have

$$\max\left\{\frac{1+n}{1+a}, \frac{1}{1+a}\right\} < \frac{1+n}{2+n}.$$

Choose $\alpha \in \mathbb{R}$ such that

$$\max\left\{\frac{1+n}{1+a}, \frac{1}{1+a}\right\} < \alpha < \frac{1+n}{2+n}. \tag{2.19}$$

Define $\bar{u} := \alpha(1 + a)$. Notice that $\bar{u} = U(\alpha, \alpha)$. It follows from (2.19) that the indifference curve passing through (α, α) intersects with the line

$$C^1 + \frac{C^2}{1+n} = 1$$

at two distinct points (A and B in Figure 2.2).

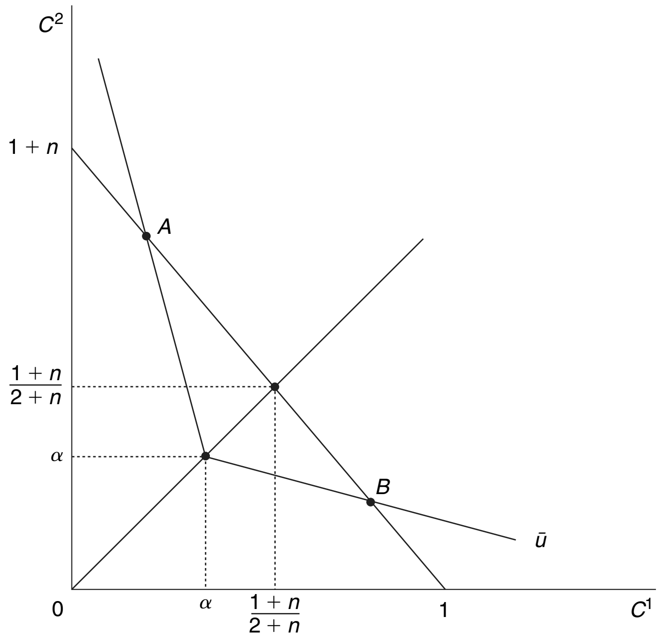

Figure 2.2 Piecewise-linear utility function

Let Z^- be the set of all negative integers, and Z^+ be the set of all positive integers. Consider the following system of equations:

$$\begin{cases} aC_t^1 + C_t^2 = \bar{u} & \text{for all } t \in Z^- \\ C_t^1 + \frac{C_{t-1}^2}{1+n} = 1 & \text{for all } t \in \{0\} \cup Z^- \\ C_0^1 = \alpha \end{cases}$$

The solution for this system of equations is:

$$C_t^1 = [\bar{u} - (1+n)]\left[\frac{1}{a} + \frac{1+n}{a^2} + \cdots + \frac{(1+n)^{-t-1}}{a^{-t}}\right] + \frac{(1+n)^{-t}\alpha}{a^{-t}} \quad (2.20)$$

$$C_t^2 = (1+n) - [\bar{u} - (1+n)]\left[\frac{1+n}{a} + \frac{(1+n)^2}{a^2} + \cdots + \frac{(1+n)^{-t-1}}{a^{-t-1}}\right] - \frac{(1+n)^{-t}\alpha}{a^{-t-1}} \quad (2.21)$$

for all $t \in Z^-$, and $C_0^1 = \alpha$. Because $a > 1+n$, it follows that

$$\lim_{t\to-\infty} C_t^1 = \frac{\bar{u}-(1+n)}{a-(1+n)}$$
$$\lim_{t\to-\infty} C_t^2 = (1+n)\left[1-\frac{\bar{u}-(1+n)}{a-(1+n)}\right].$$

Let $\bar{C}^1 := \frac{\bar{u}-(1+n)}{a-(1+n)}$ and $\bar{C}^2 := (1+n)\left[1-\frac{\bar{u}-(1+n)}{a-(1+n)}\right]$. Notice that $\bar{C}^1 + \frac{\bar{C}^2}{1+n} = 1$, and $(\bar{C}^1, \bar{C}^2) = A$ in Figure 2.2.

Next, consider the following system of equations:

$$\begin{cases} C_t^1 + aC_t^2 = \bar{u} & \text{for all } t \in Z^+ \\ C_t^1 + \frac{C_{t-1}^2}{1+n} = 1 & \text{for all } t \in Z^+ \\ C_0^2 = \alpha \end{cases}$$

The solution for this system of equations is:

$$C_t^1 = 1-(\bar{u}-1)\left[\frac{1}{a(1+n)}+\frac{1}{a^2(1+n)^2}+\cdots+\frac{1}{a^{t-1}(1+n)^{t-1}}\right] - \frac{\alpha}{a^{t-1}(1+n)^t} \tag{2.22}$$

$$C_t^2 = (\bar{u}-1)\left[\frac{1}{a}+\frac{1}{a^2(1+n)}+\cdots+\frac{1}{a^t(1+n)^{t-1}}\right]+\frac{\alpha}{a^t(1+n)^t} \tag{2.23}$$

for all $t \in Z^+$, and $C_0^2 = \alpha$. Since $a > \frac{1}{1+n}$, we have

$$\lim_{t\to\infty} C_t^1 = 1-\frac{\bar{u}-1}{a(1+n)-1}$$
$$\lim_{t\to\infty} C_t^2 = \frac{(1+n)(\bar{u}-1)}{a(1+n)-1}.$$

Let $\bar{\bar{C}}^1 := 1-\frac{\bar{u}-1}{a(1+n)-1}$ and $\bar{\bar{C}}^2 := \frac{(1+n)(\bar{u}-1)}{a(1+n)-1}$. Note that $\bar{\bar{C}}^1 + \frac{\bar{\bar{C}}^2}{1+n} = 1$, and $(\bar{\bar{C}}^1, \bar{\bar{C}}^2) = B$ in Figure 2.2.

The allocation $\{\mathbf{C}_t\}_{t\in Z}$ defined by (2.20), (2.21), (2.22), (2.23) and $(C_0^1, C_0^2) = (\alpha, \alpha)$ is exactly feasible and satisfies NELC (that is, $U(\mathbf{C}_t) = U(\mathbf{C}_{t'})$ for all $t, t' \in Z$), but it is not stationary. It should be noted, however, that $\mathbf{C}_t$ converges to a consumption bundle at a stationary allocation when $t \to -\infty$, and it also converges to a consumption bundle at another stationary allocation when $t \to +\infty$.

The crucial point in the above example is that there are indifference curves cutting the line $C^1 + \frac{C^2}{1+n} = 1$ twice. Then, we can construct a non-stationary allocation such that the consumption bundle converges to one intersection as t goes to infinity, and to the other intersection as t goes to minus infinity.

In contrast, when the utility function is linear or quasi-linear with $v'(1) > 1+n$ or $v'(0) < 1+n$, then every indifference curve cuts the line $C^1 + \frac{C^2}{1+n} = 1$ at most once. In this case, any non-stationary allocation satisfying NELC must violate feasibility either as t goes to infinity, or as t goes to minus infinity.

2.6 Pareto efficiency and maximal utility under NELC

In this section, we only assume that $U(C^1, C^2)$ is continuous. Within this most general framework, we settle several problems on the existence of relevant allocations with the purpose of securing the non-emptiness of our analysis of the properties of these allocations.

Proposition 6 *There exists a Pareto efficient allocation.*

Proof: Let $\{\gamma_t\}_{t\in Z}$ be a doubly infinite sequence with $\gamma_t > 0$ for all t and $\sum_{t\in Z} \gamma_t = 1$. Let F be the set of exactly feasible allocations. Clearly, F is non-empty and compact in the product topology. For each $\mathbf{C} \in F$, let $V(\mathbf{C}) = \sum_{t\in Z} \gamma_t U(C_t^1, C_t^2)$. It is easy to see that V is product continuous on F. By the Weierstrass theorem, there exists $\mathbf{C}^*$ in F that attains the maximum value of V over F. Clearly, $\mathbf{C}^*$ is Pareto efficient. *Q.E.D.*

Next, we consider the following procedure to select allocations. First, we choose all exactly feasible allocations satisfying NELC. Then, we select 'optimal' allocations, from an efficiency standpoint, among those allocations. The selection procedure is based on the *equity-first and efficiency-second principle* due to Tadenuma (2002).

Let *NE* be the set of all exactly feasible allocations satisfying NELC. Clearly, *NE* is non-empty. Recall that at any allocation satisfying NELC, all the generations t attain the same level of utility. Let

$$u^* := \sup \{\alpha \mid \exists \mathrm{C} \in NE \text{ such that } \forall t \in Z, U(C_t^1, C_t^2) = \alpha\}.$$

We say that an allocation $\mathbf{C} \in NE$ *attains maximal utility under NELC* if $U(C_t^1, C_t^2) = u^*$ for all $t \in Z$.

Proposition 7 *There exists an allocation that attains maximal utility under NELC.*

Proof: Let $\{\{\mathbf{C}_t^n\}_{t\in Z}\}_{n=1}^{\infty}$ be a sequence in *NE* such that $\{U(C_t^{1n}, C_t^{2n})\}_{n=1}^{\infty}$ converges to u^*. Since $\{\mathbf{C}_t^n\}_{t\in Z}$ is exactly feasible, there exists a subsequence $\{\{\mathbf{C}_t^{n_q}\}_{t\in Z}\}_{q=1}^{\infty}$ such that for each $t\{\mathbf{C}_t^{n_q}\}_{q=1}^{\infty}$ converges to some $\mathbf{C}_t^*$. Clearly, $\{\mathbf{C}_t^{n*}\}_{t\in Z}$ is exactly feasible. Since $U(C^1, C^2)$ is continuous, $U(C_t^{1*}, C_t^{2*}) = \alpha$. Thus, $\{\mathbf{C}_t^{n*}\}_{t\in Z}$ attains maximal utility under NELC. *Q.E.D.*

Remark: It is an open question whether an allocation that attains maximal utility under NELC is Pareto efficient or not.

If the utility function U is, in addition, quasi-concave, one can prove a stronger result.

Proposition 8 *If U is, in addition, quasi-concave, then there exists a stationary allocation that attains maximal utility under NELC.*

Proof: Let $\{\mathbf{C}_t^{n*}\}_{t\in Z}$ be an allocation that attains maximal utility under NELC. By averaging operations, we construct a sequence of allocations that attain maximal utility under NELC. For each positive integer N and for each $t \in Z$, let

$$\mathbf{C}_t^N = \frac{1}{2N+1}(\mathbf{C}_{t-N}^* + \mathbf{C}_{t-N+1}^* + \cdots + \mathbf{C}_{t-1}^* + \mathbf{C}_t^* + \mathbf{C}_{t+1}^* + \cdots + \mathbf{C}_{t+N}^*).$$

Then, it is easy to see that $\mathbf{C}^N = \{\mathbf{C}_t^N\}_{t\in Z}$ is exactly feasible. By quasi-concavity of U,

$$U(\mathbf{C}_t^N) \geq \min_{t-N\leq n\leq t+N}\{U(\mathbf{C}_n^N)\} = u^*$$

for each $t \in Z$, where the last equality follows from the hypothesis that $\{\mathbf{C}_t^{n*}\}_{t\in Z}$ attains maximal utility under NELC. By exact feasibility, the sequence $\{\mathbf{C}_t^N\}_{N=1}^{\infty}$ is uniformly bounded for each $t \in Z$. By Cantor's diagonalization process, there exists a subsequence $\{\mathbf{C}^{N_q}\}_{q=1}^{\infty} = \{\{\mathbf{C}^{N_q}\}_{t\in Z}\}_{q=1}^{\infty}$ such that for each $t \in Z$, $\{\mathbf{C}^{N_q}\}_{q=1}^{\infty}$ converges to some $\mathbf{C}_t$. Clearly, $\mathbf{C} = \{\mathbf{C}_t\}_{t\in Z}$ is exactly feasible. By continuity of U, $U(\mathbf{C}_t) \geq u^*$ for each $t \in Z$.

Now, we will show that $\mathbf{C}_t = \mathbf{C}_{t-1}$ for each $t \in Z$. Suppose that $\mathbf{C}_t \neq \mathbf{C}_{t-1}$ for some $t \in Z$. Let $\delta = |\mathbf{C}_t - \mathbf{C}_{t-1}|$. Then, $\delta > 0$. For each $t \in Z$,

$$|\mathbf{C}_t^N - \mathbf{C}_{t-1}^N| = \frac{|\mathbf{C}_{t+N}^* - \mathbf{C}_{t-1-N}^*|}{2N+1} \leq \frac{2}{2N+1}.$$

Hence, there exists a positive integer $N(\delta)$ such that for all $N \geq N(\delta)$, $|\mathbf{C}_t^N - \mathbf{C}_{t-1}^N| < \frac{\delta}{2}$. On the other hand, we have

$$\begin{aligned}\delta = |\mathbf{C}_t - \mathbf{C}_{t-1}| \leq |\mathbf{C}_t - \mathbf{C}_t^N| + |\mathbf{C}_t^N - \mathbf{C}_{t-1}^N| + |\mathbf{C}_{t-1}^N - \mathbf{C}_{t-1}| \\ < \frac{\delta}{4} + |\mathbf{C}_t^N - \mathbf{C}_{t-1}^N| + \frac{\delta}{4}\end{aligned}$$

holds for all $N \geq N(\delta)$. Hence, for all $N \geq N(\delta)$,

$$\frac{\delta}{2} < |\mathbf{C}_t^N - \mathbf{C}_{t-1}^N|.$$

This is a contradiction. Thus, $\mathbf{C}_t = \mathbf{C}_{t-1}$ for each $t \in Z$. Let us denote this common vector by $\tilde{\mathbf{C}} = (\tilde{C}^1, \tilde{C}^2)$. By exact feasibility of $\mathbf{C} = \{\mathbf{C}_t\}_{t \in Z}$,

$$\tilde{C}^1 + \frac{\tilde{C}^2}{1+n} = 1.$$

Clearly, $U(\tilde{\mathbf{C}}) \geq u^*$. By maximality under NELC, $U(\tilde{\mathbf{C}}) = u^*$. Therefore, $\tilde{\mathbf{C}} = (\tilde{C}^1, \tilde{C}^2)$ generates a stationary allocation that attains maximal utility under NELC. *Q.E.D.*

Corollary 4 *A consumption vector* $\hat{\mathbf{C}} = (\hat{C}^1, \hat{C}^2)$ *generates a stationary allocation that attains maximal utility under NELC if and only if it maximizes* $U(\mathbf{C})$ *subject to* $C^1 + C^2/(1+n) = 1$.

Proof: Let $\tilde{\mathbf{C}} = (\tilde{C}^1, \tilde{C}^2)$ be a consumption vector generating a stationary allocation that attains maximal utility under NELC. By maximality under NELC, for any consumption vector $\hat{\mathbf{C}} = (\hat{C}^1, \hat{C}^2)$ satisfying $\hat{C}^1 + \hat{C}^2/(1+n) = 1$, $U(\tilde{\mathbf{C}}) \geq U(\hat{\mathbf{C}})$.

Conversely, let $\hat{\mathbf{C}}$ be a consumption vector maximizing $U(\mathbf{C})$ subject to $\hat{C}^1 + \hat{C}^2/(1+n) = 1$. Then, $U(\hat{\mathbf{C}}) \geq U(\tilde{\mathbf{C}})$. Since $\tilde{\mathbf{C}}$ attains maximal utility under NELC, $U(\tilde{\mathbf{C}}) \geq U(\hat{\mathbf{C}})$. Hence, $U(\tilde{\mathbf{C}}) = U(\hat{\mathbf{C}})$. This completes the proof. *Q.E.D.*

2.7 Conclusion

The main conclusions of our analysis may be summarized in Figure 2.3, where an arrow represents a logical implication which cannot be reversed in general.

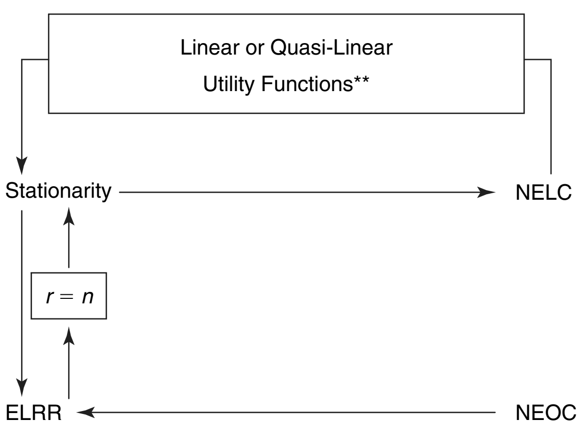

Figure 2.3 Logical relations among equity concepts*
* Exact feasibility is always assumed.
** For the case of quasi-linear utility functions, we need an assumption that $\min\{v'(c) \mid c \in [0,1]\} > 1+n$ or $\max\{v'(c) \mid c \in [0,1]\} < 1+n$.

An arrow from NELC to stationarity is valid subject to the condition that the utility function is either linear or quasi-linear.

Recollect that our three alternative concepts of intergenerational equity are not just hairsplitting theoretical toys, but are relevant in many actual disputes in the design and implementation of equitable social security schemes and intergenerational transfer mechanisms. Despite their obvious importance, not many studies have been done on equity and efficiency in overlapping generations economies. It is hoped that our exploration of the fundamental implications of several distinct notions of equity-as-no-envy may be found useful in orienting further studies of this important issue.

References

Cass, D. and M. E. Yaari (1966) 'Re-Examination of the Pure Consumption Loans Model', *Journal of Political Economy*, vol. 74, pp. 353–67.

Foley, D. (1967) 'Resource Allocation and the Public Sector', *Yale Economic Essays*, vol. 7, pp. 45–98.

Gale, D. (1973) 'Pure Exchange Equilibrium of Dynamic Economic Models', *Journal of Economic Theory*, vol. 6, pp. 12–36.

Kolm, S.-C. (1972) *Justice et Equité* (Paris: Ed. du CNRS).

Samuelson, P. A. (1958) 'An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money', *Journal of Political Economy*, vol. 66, pp. 467–82.

Suzumura, K. (2002) 'On the Concept of Intergenerational Equity', mimeo.

Tadenuma, K. (2002) 'Efficiency First or Equity First? Two Principles and Rationality of Social Choice', *Journal of Economic Theory*, vol. 104, pp. 462–72.

3

Social Security Pensions and Intergenerational Equity: The Japanese Case*

Noriyuki Takayama

3.1 Introduction

Japan already has the oldest population in the world. It built a generous social security pension programme, but since 2001 the income surplus of the principal pension programme has turned into a deficit, and from then until the 2004 reforms, its balance sheet, which showed a huge excess of liabilities, engendered a growing distrust of the government's commitment on pensions. The Japanese have been increasingly concerned with the incentive-compatibility problem.

Social security pensions are a contract between the government and an individual. For any individual, the contract is to continue to receive a pension after retirement at 60 years or more, and more broadly it is a contract between generations, relating to an uncertain future, adaptation to which requires timely and proper revisions.

This chapter addresses equity issues between generations in the context of social security pensions in Japan, with special attention being paid to two problems. The first is how to find an intergenerationally-equitable remedy for any past mistakes and the second is to devise a pension scheme avoiding any inequities between generations that arise from prospective uncertainties.

Before entering that discussion, the chapter briefly sketches the Japanese social security pension programme and summarizes its major problems in the context of the 2004 pension reforms.

3.2 Brief outline of pension provisions before the 2004 Reform

Since 1980, Japan has repeated piecemeal pension reforms every five years, mainly under the impact of anticipated demographic and economic factors.

* The author is greatly indebted to the participants in the Hakone Conference, particularly Professors T. Ihori, S. Ogura, T. Oshio and J. Silvestre, for valuable comments on an earlier draft.

Since that date the excessively generous pension benefits have been reduced step by step with an increase of the normal pensionable age from 60 to 65, with the pension contribution rate also gradually lifted. Nevertheless, pension provisions remain generous and generate serious financial difficulties in the future.

Japan currently has a two-tier benefit system, providing all groups of the population, the first-tier, flat-rate basic benefit. The second-tier, earnings-related benefit, applies only to employees.[1] The system operates largely as a pay-as-you-go defined benefit programme.

The flat-rate basic pension covers all residents aged 20 to 60. The full old-age pension is payable after 40 years of contributions, provided the contributions were made before 60 years of age. The maximum *monthly* pension of 66,200 yen at 2004 prices (with maximum number of years of coverage) per person is payable from age 65.[2] The benefit is indexed automatically each fiscal year (from 1 April) to reflect changes in the consumer price index (CPI) from the previous calendar year. The pension may be claimed at any age between 60 and 70 years. It is subject to actuarial reduction if claimed before age 65, or actuarial increase if claimed after 65 years.

Earnings-related benefits are given to all employees. The accrual rate for the earnings-related component of old-age benefits is 0.5481 per cent per year: 40 years' contributions will thus earn 28.5 per cent of career average monthly real earnings.[3]

The career-average monthly earnings are calculated over the employee's entire period of coverage, adjusted by a net-wage index factor, and converted to the current earnings level. The full earnings-related pension is normally payable from age 65 to an employee who is fully retired.[4] An earnings test is applied to those who are not fully retired. The current replacement rate (including basic benefits) for take-home pay or net income is about 60 per cent for a 'model' male retiree (with an average salary earned during 40 years of coverage) and his dependent wife. Its *monthly* benefit is about 233,000 yen.

Equal percentage contributions are required of employees and their employers, based on the annual standard earnings, including bonuses. The total percentage in effect before October 2004 was 13.58 per cent for the principal programme for private-sector employees (Kosei-Nenkin-Hoken, KNH). Non-employed persons between the age of 20 to 60 years pay flat-rate individual contributions. The current rate since April 1998 is 13,300 yen per month. For those who cannot pay for financial reasons, exemptions are permitted: the flat-rate basic benefits for the period of exemption are one-third of the normal amount.

Under the current system, if the husband has the pension contribution for social security deducted from his salary, his dependent wife is automatically entitled to the flat-rate basic benefits, and she is not required to make any individual payments to the public pension system. The government subsidizes one-third of the total cost of the flat-rate basic benefits. There is no subsidy for the earnings-related part. The government also pays administrative expenses.

The aggregate amount of social security pension benefits was around 46 trillion yen in 2004, which is equivalent to about 9 per cent of that year's GDP.

3.3 The impact of demographic factors on social security finance

The population projections of January 2002 of the Japanese National Institute of Population and Social Security Research (NIPSSR) indicate the total peaking at 128 million around 2006 and then beginning to fall steadily to about 50 per cent of the current number by 2100.

The total fertility rate (TFR) was 1.29 in 2004 and there is as yet little sign that it will stabilize or return to a higher level. Yet the 2002 *medium variant* projections assume that after a historical low of 1.31 in 2006 it will gradually rise to 1.39 around 2050, progressing slowly to 2.07 by 2150. The number of births, about 1.12 million in 2003, will continue to decrease to less than 1.0 million by 2014, falling further to 0.67 million in 2050.

Because it has the longest life expectancy, Japan is now experiencing a very rapid ageing of its population. The number of the elderly (65 years and above) stood at 24.9 million in 2004, and is forecast to increase sharply to reach 34 million by 2018, and then remain around 34–36 million until around 2060. Consequently, the proportion of the elderly will rise rapidly from 19.5 per cent in 2004 to 25.3 per cent by 2014, rising further to more than 30 per cent by 2033. Japan already has one of the oldest populations in the world.

In Japan, nearly 70 per cent of social security benefits are currently distributed to the elderly. Along with the ailing domestic economy, the rapid aging will certainly put more and more stresses on financing social security.

In May 2004, the Ministry of Health, Labour and Welfare of Japan published estimates of the cost of social security, using the 2002 population projections of the NIPSSR. Based on these estimates, the aggregate cost of social security was 17.2 per cent of GDP in 2004 and will steadily increase to 24.3 per cent by 2025, if the current provisions for benefits remain unchanged.

Of the various costs, that of pensions is quite predominant, amounting to 9 per cent of GDP in 2004, with further projected increases to 11.6 per cent by 2025. The cost for health care is 5.2 per cent in 2004, but will rapidly rise to 8.1 per cent by 2025.

The Japanese economy is still reeling from the effects of its burst bubble, and the decline in population will soon be reflected in a sharp decline in young labour, in a falling savings rate and in a decrease in capital formation, all of which will contribute to a further shrinking of the country's economy.

3.4 Some basic facts on pensions

Social security pensions in Japan currently face several difficulties. Among others, the following five are particularly pressing.

Persistent deficit in the income statement

Since 2001, the KNH has been facing an income statement deficit. It recorded a deficit of 700 billion yen in 2001, and the deficit would be 4.2 trillion yen in 2002. It is estimated that the deficit will persist for a long time, unless radical remedies are made in the KNH financing.

Huge excess liabilities in the balance sheet

The KNH balance sheet is shown in Figure 3.1. In calculating the balance sheet, we assumed that:

(i) annual increases in wages and CPI are 2.1 per cent and 1.0 per cent respectively in nominal terms, while the discount rate is 3.2 per cent annually,

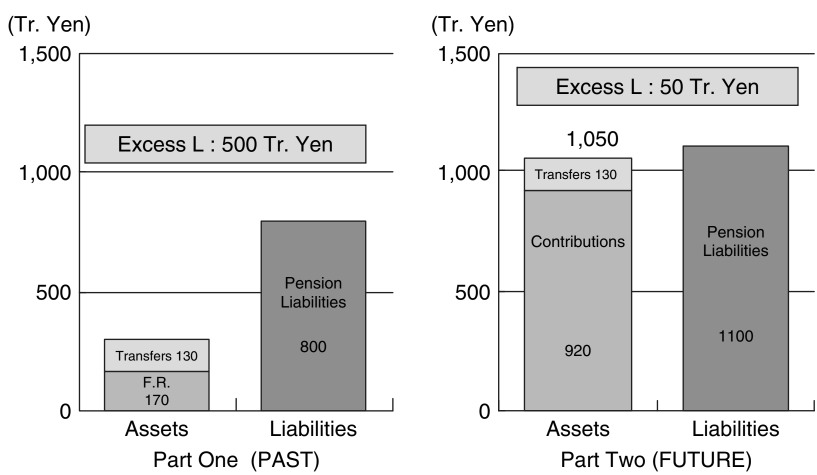

Figure 3.1 KNH balance sheet: before reforms (as at 31 March 2005)

(ii) the current contribution rate of the KNH, 13.58 percentage point, will remain unchanged in the future, and
(iii) the period up to year 2100 is taken into account.

Figure 3.1 indicates that as at 31 March 2005, there would be excess liabilities of 550 trillion yen, which is a quarter of the total liabilities.[5]

Part One of Figure 3.1 is assets and liabilities accrued from past contributions and Part Two is those accrued from future contributions. Figure 3.1 implies that, as far as Part Two is concerned, that balance sheet of the KNH has been almost cleaned up. The funding sources of the current provisions will be sufficient to finance future benefits, and the only task left is to slim down future benefits by 4.5 per cent.

But if we look at Part One of Figure 3.1, things appear quite different. The remaining pension liabilities are estimated to be 800 trillion yen, while pension assets are only 300 trillion yen (a funded reserve of 170 trillion yen plus transfers from general revenue of 130 trillion yen). The difference is quite large – about 500 trillion yen,[6] which accounts for the major part of excess liabilities in the KNH.

500 trillion yen is more than 60 per cent of Part One liabilities, equivalent to about 100 per cent of GDP of Japan in 2004. In the past, too many promises on pension benefits were made, while sufficient funding sources have not been arranged. The Japanese have enjoyed a long history of generous social security pensions. However, contributions made in the past were relatively small, resulting in a fairly small funded reserve. Consequently, the locus of the true crisis in Japanese social security pensions is how to handle the excess liabilities of 500 trillion yen which were entitled from contributions made in the past.

Pension contributions: heavy burdens outstanding

In Japanese public debates, one of the principal issues has been how to cut down personal and corporate income tax. But recently situations changed drastically. Social security contributions (for pensions, health care, unemployment, work injury and long-term care) are 55.6 trillion yen (15.2 per cent of national income) for FY 2003. This is apparently more than all tax revenues (43.9 trillion yen) of the central government for the same year. Since 1998, the central government has acquired more from social security contributions than from tax incomes. Looking at more details, we can find that revenue from personal income tax is 13.8 trillion yen and corporate income tax is 9.1 trillion yen, while revenue from social security pension contributions stands out – at 29.0 trillion yen. Needless to say, the last apparently places a most heavy burden on the public. The Japanese now feel that social security pension contributions are too heavy; they operate as the most significant factor in determining the take-home pay from the gross salary.

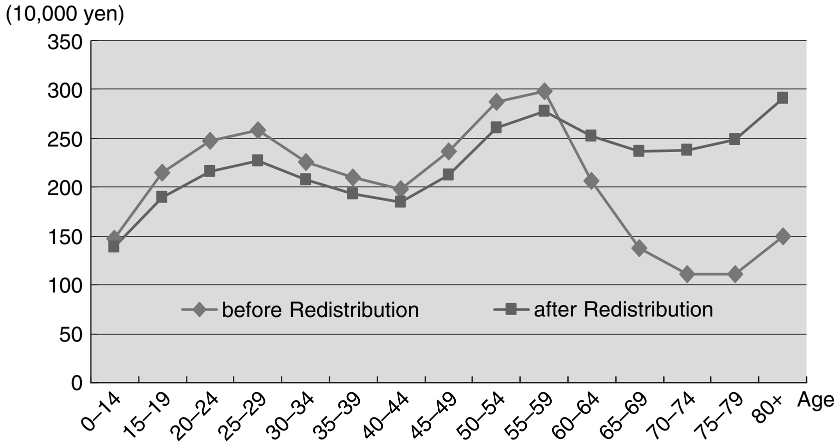

Figure 3.2 Per-capita income by age in Japan
Source: Ministry of Health and Welfare, *The 1996 Income Redistribution Survey*.

On the other hand, managements have begun to show serious concerns on any further increases in social security contributions.

Overshooting in income transfer between generations

It may be amazing that currently in Japan the elderly are better-off than those aged 30 to 44 in terms of per-capita income after redistribution (see Figure 3.2). Undoubtedly, there must still be room for reduction in benefits provided to the current retired population.

Increasing rate of drop-out

In the past twenty years, the Japanese government has made repeated changes to the pension programme, increasing social security pension contributions and reducing benefits through raising the normal pensionable age while reducing the accrual rate. Similar piecemeal reforms are likely to follow in the future.

Many Japanese feel that the government is breaking its promise. As distrust against government commitment builds up, concern on such an 'incredibility problem' is also growing.

In 2002 nearly 50 per cent of non-salaried workers and persons with no occupations dropped out from the basic level of old-age income protection, owing to exemption, delinquency in paying contributions or non-application (see Figure 3.3 for increasing delinquency).

Also, employers are carefully trying to find ways of avoiding to pay social security pension contributions. Indeed, the aggregate amount of the KNH

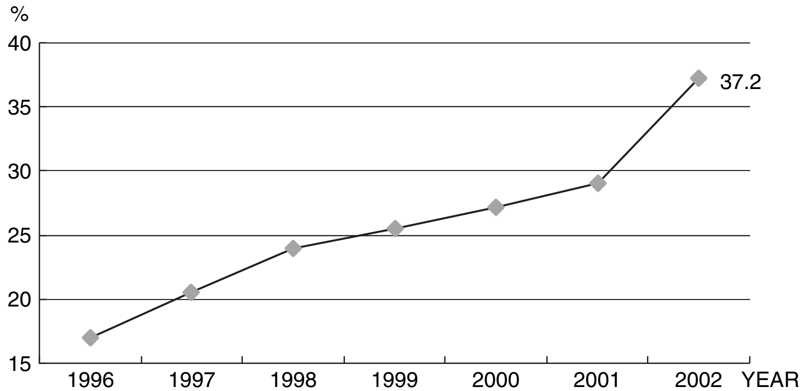

Figure 3.3 Drop-out from Social Security Pensions (Non-employees): Delinquency in Paying Pension Contributions

contributions has been decreasing since 1998, in spite of no change in the contribution rate.

Any further escalation in the social security contribution rate will surely induce a higher drop-out rate.[7]

3.5 The 2004 pension reform: main contents and remaining difficulties[8]

On 10 February 2004, the administration of Prime Minister Koizumi Jun'ichirō submitted a set of pension reform bills to the National Diet, which were enacted on 5 June. This section describes the gist of the approved reforms and explores issues that remain to be addressed.

Increases in contributions

Salaried workers are, as a rule, enrolled in the KNH, which is part of the public pension system. Since October 1996 contributions under this plan have been set at 13.58 per cent of annual income, paid half by the worker and half by the employer, but the newly enacted reforms will raise this rate by 0.354 percentage points per year starting in October 2004. The rate will rise every September thereafter until 2017, after which it will remain fixed at 18.30 per cent. The portion paid by workers will accordingly rise from the current 6.79 per cent of annual income to 9.15 per cent.

For an 'average' male company employee earning 360,000 yen a month plus annual bonuses equivalent to 3.6 months' pay, contributions will increase by nearly 20,000 yen a year starting this October 2004, and by the time they stop rising in September 2017, they will have reached just under 1.03 million yen a year, and the share paid by the worker will be just over 514,000 yen. This comes to 35 per cent more than the current level of contributions.

Those who are not enrolled in the KNH or other public pension schemes for civil servants are required to participate in the National Pension plan, which provides just the so-called basic pension (the basic pension also forms the first tier of benefits under the KNH and other public pension systems for civil servants). Contributions under this plan will rise by 280 yen each April from the current 13,300 yen per month until they plateau at 16,900 yen (at 2004 prices) in April 2017. The actual rise in National Pension contribution will be adjusted according to increases in general wage levels.

In addition, the government will increase its subsidies for the basic pension. One-third of the cost of basic pension benefits is paid from the national treasury; this share is to be raised in stages until it reaches one-half in 2009.

Reductions in benefits

Benefits under the KNH consist of two tiers: the flat-rate basic pension, which is paid to all public pension plan participants, and a separate earnings-related component. The latter is calculated on the basis of the worker's average pre-retirement income, converted to current values. Until now, the index used to convert past income to current values was the rate of increase in take-home pay. Under the recently enacted reforms, though, this index will be subject to a negative adjustment over the course of an 'exceptional period' based on changes in two demographic factors, namely, the decline in the number of participants and the increase in life expectancy. This period of adjustment is expected to last through 2023.

The application of the first demographic factor will mean that benefit levels will be cut to reflect the fact that fewer people are supporting the pension system. The actual number of people enrolled in all public pension schemes will be ascertained each year, and the rate of decline will be calculated based on this figure. The average annual decline is projected to be around 0.6 points.

Introducing the second demographic factor, meanwhile, will adjust for the fact that people are living longer and thus collecting their pensions for more years; the aim is to slow the pace of increase in the total amount of benefits paid as a result of increased longevity. This factor will not be calculated by tracking future movements in life expectancy; instead, it has been set at an annual rate of about 0.3 percentage points on the basis of current demographic projections for the period through 2025. Together, the two demographic factors are thus expected to mean a negative adjustment of about 0.9 points a year during the period in question.

How will these changes affect people's benefits in concrete terms? Let us consider the case of a pair of 'model' KNH beneficiaries as defined by the Ministry of Health, Labor, and Welfare: a 65-year-old man who earned the average wage throughout his 40-year career and his 65-year-old wife who was a full-time homemaker for 40 years from her twentieth birthday. In fiscal 2004

(April 2004 to March 2005), this model couple would receive 233,000 yen a month.

How does this amount compare to what employees are currently taking home? The average monthly income of a salaried worker in 2004 is projected to be around 360,000 yen, before taxes and social insurance deductions. Assuming that this is supplemented by bonuses totalling an equivalent of 3.6 months' pay, the average annual income is roughly 5.6 million yen. Deducting 16 per cent of this figure for taxes and social insurance payments leaves a figure for annual take-home pay of about 4.7 million yen, or 393,000 yen a month.

The 233,000 yen provided to the model pensioners is approximately 59.3 per cent of 393,000 yen. But this percentage, which pension specialists call the 'income replacement ratio,' will gradually decline to an estimated figure of 50.2 per cent as of fiscal 2023 (assuming that consumer prices and nominal wages rise according to government projections by 1.0 per cent and 2.1 per cent a year, respectively). Over the next two decades, then, benefit levels will decline by roughly 15 per cent by comparison with wage levels.

The revised pension legislation stipulates that the income replacement ratio is not to fall below 50 per cent for the model case described above, and so the exceptional period of negative adjustment will come to an end once the ratio declines to 50 per cent. This provision was included to alleviate fears that benefits would continue to shrink without limit.

How will the reforms affect those who are already receiving their pensions? Until now, benefits for those 65 years old and over were adjusted for fluctuations in the consumer price index. This ensured that pensioners' real purchasing power remained unchanged and helped ease post-retirement worries. But this cost-of-living link will effectively be severed during the exceptional period, since the application of the demographic factors will pull down real benefits by around 0.9 points a year. In principle, however, nominal benefits are not to be cut unless there has also been a drop in consumer prices. Once the exceptional period is over, the link to the consumer price index is to be restored.

Incentive-compatible?

Social insurance contributions in Japan already exceed the amount collected in national taxes, and contributions to the pension system are by far the biggest social insurance item. If this already huge sum is increased by more than 1 trillion yen a year, as the government plans, both individuals and companies are bound to change their behaviour. Government projections of revenues and expenditures, though, completely ignore the prospect of such changes.

Companies will likely revamp their hiring plans and wage scales to sidestep the higher social insurance burden. They will cut back on recruitment of new graduates and become more selective about mid-career hiring as well. Many young people will be stripped of employment opportunities and driven out of the labour market, instead of being enlisted to support the pension system with a percentage of their income. And most of the employment options for middle-aged women who wish to reenter the workforce will be low-paying ones. Only a few older workers will be able to continue commanding high wages; there is likely to be a dramatic rise in the number of aging workers who will be forced to choose between remaining on the payroll with a cut in pay or settling for retirement. Many more companies will either choose or be forced to leave the KNH, causing the number of subscribers to fall far below the government's projections and pushing the system closer to bankruptcy.

The jobless rate on the whole will rise. The Japan Ministry of Economy, Trade, and Industry has estimated that higher pension contributions will lead to the loss of 1 million jobs and boost the unemployment rate by 1.3 points. The government plan to increase pension contributions annually for the next 13 years will exert ongoing deflationary pressure on the Japanese economy. For the worker, a rise in contribution levels means less take-home pay; as a result, consumer spending is likely to fall, and this will surely hinder prospects for a self-sustaining recovery and return to steady growth.

Another problem with increasing pension contributions is that they are regressive, since there is a ceiling for the earnings on which payment calculations are based and unearned income is not included in the calculations at all.

One major objective of the reforms is to eventually eliminate the huge excess liabilities in the balance sheet of the KNH. The plan is to generate a surplus by (1) hiking contributions, (2) increasing payments from the national treasury, and (3) reducing benefits. The policy measures adopted in the 2004 pension reform bill will induce huge *excess assets* of 420 trillion yen in Part Two balance sheet whereby offsetting excess liabilities of the same amount in Part One balance sheet as shown in Figure 3.4. Huge excess assets of Part Two balance sheet imply that future generations will be forced to pay more than the anticipated benefits they will receive. Their benefits will be around 80 per cent of their contributions, on the whole.

It seems as if we cut paper not with scissors but with a saw. Younger generations are most likely to increase their distrust of government. The incentive-compatibility problem or the drop-out problem will become more severe. The management (Nippon Keidanren) and trade unions (Rengo) both oppose any further increases of more than 15 percentage point in the KNH contribution rate.

Declining replacement rate

As noted above, those who are already receiving their pensions will see their benefits decline in real terms by an average of 0.9 per cent per year. The government scenario sees consumer prices eventually rising by 1.0 per cent a year and take-home pay by 2.1 per cent a year. This means that the model beneficiary who begins receiving 233,000 yen a month at age 65 in 2004 will

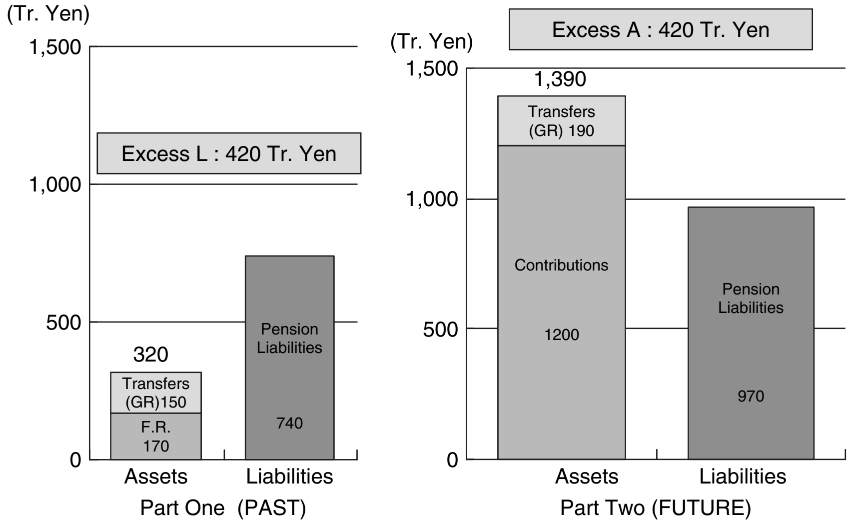

Figure 3.4 KNH balance sheet: after reforms (as at 31 March 2005)

get roughly 240,000 yen at age 84 in 2023; nominal benefits, in other words, will remain virtually unchanged for two decades, despite the fact that the average take-home pay of the working population will have risen by over 40 per cent. The income replacement rate, which stood at nearly 60 per cent at age 65, will dwindle to 43 per cent by the time the model recipient turns 84. The promise of benefits in excess of 50 per cent of take-home pay does not apply, therefore, to those who are already on old-age pensions.

Automatic balance mechanism: still incomplete

The so-called demographic factors are likely to continue changing for the foreseeable future. The government itself foresees the number of participants in public pension plans declining over the coming century: the estimated figure of 69.4 million participants as of 2005 is expected to fall to 61.0 million in 2025, 45.3 million in 2050, and 29.2 million in 2100. This corresponds to an average annual decline of 0.6 per cent through 2025, 1.2 per cent for the quarter century from 2025, and 0.9 per cent for the half century from 2050. In other words, the decline in the number of workers who are financially supporting the public pension system is not likely to stop after just two decades.

The recently enacted reforms, though, adjust benefit levels in keeping with the decline in the contribution paying population for the next 20 years only; the government's 'standard case' does not foresee any further downward revisions, even if the number of participants continues to fall. If the government really anticipates an ongoing decline, there is no good reason to abruptly stop adjusting benefit levels after a certain period of time. Sweden and Germany, for instance, have adopted permanent mechanisms whereby benefit levels are automatically adjusted for fluctuations in demographic factors.

The decision to keep the model income replacement rate at 50 per cent at the point when pension payments commence represents, in effect, the adoption of a defined benefit formula. Maintaining both fixed contributions on the one hand and defined benefit levels on the other is not an easy task, for there is no room to deal flexibly with unforeseen developments. The government will be confronted with a fiscal emergency should its projections for growth in contributions and a reversal in the falling birthrate veer widely from the mark.

The government bases its population figures on the January 2002 projections of the National Institute of Population and Social Security Research. Under these projections, the medium variant for the total fertility rate (the average number of childbirths per woman) falls to 1.31 in 2007, after which it begins climbing, reaching 1.39 in 2050 and 1.73 in 2100. Actual figures since the projections were released have been slightly lower than this variant, and there are no signs whatsoever that the fertility rate will stop declining in 2007.

The normal pensionable age

If the government is to keep its promise about an upper limit for contributions and a lower limit for benefits, the only policy option it will have in the event of a financial shortfall will be to raise the age at which people begin to receive benefits. The reform package makes no mention of such a possibility; the drafters of the bills no doubt chose to simply put this task off to a future date.[9]

Why increasing transfers from general revenue?

By fiscal year 2009 the share of the basic pension benefits funded by the national treasury will be raised from one-third to one-half. This means that more taxes will be used to cover the cost of benefits. Taxes are by nature different from contributions paid by participants in specific pension plans, and there is a need to reconsider the benefits that are to be funded by tax revenues.

The leaders of Japanese industry tend to be quite advanced in years. For the most part, they are over the age of 65, which means that they are qualified to receive the flat-rate basic pension. Even though they are among the wealthiest people in the country, they are entitled to the same basic pension as other older people hovering around the poverty line. Using tax revenues to finance a bigger share of the basic pension essentially means asking taxpayers to foot a bigger bill for the benefits of wealthy households as well. For an elderly couple, the tax-financed portion of the basic pension will rise from 530,000 yen a year to 800,000 yen. If a need arises to raise taxes at a future date, who will then actually agree to pay more? Few people will be willing to tolerate such wasteful uses of tax money.

3.6 Intergenerational equity issues in Japan's social security pensions

The huge excess liabilities of 500 trillion yen appearing in Part One of Figure 3.1 partly reflect mistakes made in the past.[10] It is true that any social security scheme for pensions faces great uncertainties for its future long-term scenarios; uncertainties on the number of the participants, the number of the pensioners, the rate of increases in wages or the consumer price index, and the rate of return from investment. No one has precise information on these variables beforehand. Yet, the system planners need some fixed figures on the system's future scenario in designing (or re-designing) the pension system. It is often the case, however, that the assumed figures are more or less different from the actual ones. What really matters is whether or not the system planners adjust their system to correspond to the changing circumstances in a timely and proper way.

Japanese experiences in the past 30 years show that the adjustments were so slow and insufficient as to produce huge excess liabilities amounting to 500 trillion yen. It is evident that pension projections always turned out to be too optimistic, and that politicians were always reluctant to introduce painful remedies for current pensioners and current contributors, leaving the pension system financially unsustainable and inequitable between generations.

The 2004 reform in Japan looks very drastic, since the introduction of the demographic factors will significantly reduce the level of pension benefits in real terms. It is regarded as an inevitable rebound or a backfire to make up for the mistakes or omissions made in the past, which the current pensioners and the baby-boomers were responsible for. Yet the 2004 reform still suffers from an incentive-compatibility problem, leaving the pension system inequitable from the younger generations' point of view.[11]

3.7 Future policy options for securing equity between generations

Are there any policy measures which could avoid the incentive-compatibility problem in Japan? This section tries to answer this question.

To begin with, how about separating the 'legacy pension' problem from the problem of re-building a sustainable pension system for the future? The problems are quite different in nature and accordingly they require separate handlings.

The legacy pension problem of Japan looks like *sunk costs* in the economic perspective. It can be eased not by increasing the KNH contribution rate, but instead introducing a new 2.0 per cent earmarked consumption tax and intensive interjection of the increased transfers from general revenue (see Figure 3.5). Needless to say, the current generous benefits have to be reduced

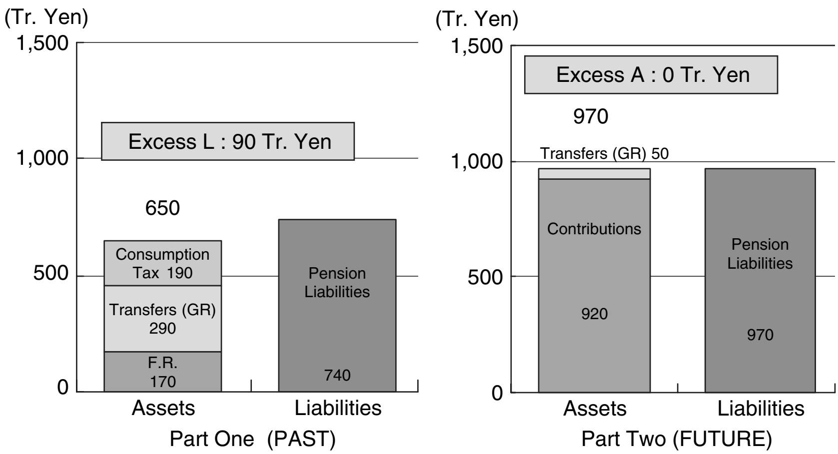

Figure 3.5 KNH balance sheet: alternative reforms (as at 31 March 2005)

more or less by the same percentage in the aggregate level as implemented in the 2004 pension reform.

All these measures are considered from the understanding that current pensioners and baby-boomers are mainly responsible for Part One excess liabilities, and thus they are to come first to diminish the existing excess liabilities. Note that any increases in the contribution rate for social security pensions will be paid by current younger and future generations. Current pensioners no longer pay them and baby-boomers will pay them only a little. They are not an appropriate measure for diminishing Part One excess liabilities. By contrast, an increase in the consumption-based tax will be shared by all the existing and future generations including current pensioners and baby-boomers.[12] Increased transfers from general revenue can be financed by increases in inheritance tax and income tax on pension benefits, as well.[13,14]

When it comes to the Part Two balance sheet, which relates to future contributions and promised pension benefits entitled by future contributions, a switch to the NDC (notional defined contribution) is possible and preferable.[15] The KNH contribution rate can be kept unchanged at the current level around 14 percentage points.

With the NDC plan, the incentive-compatibility problem can be avoided. Indeed, every penny counts in the NDC, and this would be the most important element when we switch to an NDC plan. It will be demonstrated to the public that everybody gets a pension equivalent to his/her own contribution payments (see Könberg 2002, and Palmer 2003).

In the NDC, the notional rate of return should be endogenous. It can be periodically adjusted by an automatic balance mechanism introduced in

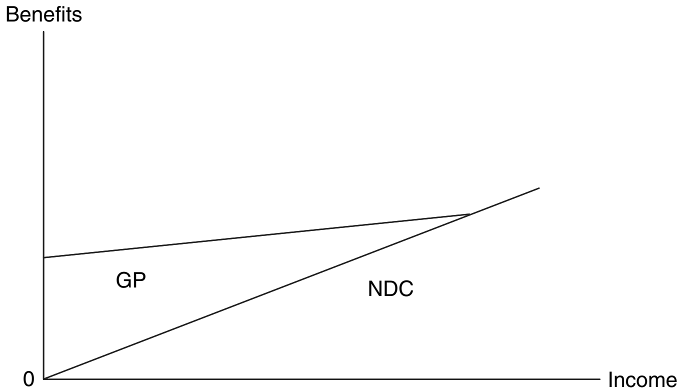

Figure 3.6 NDC plus GP

Sweden (see Settergren, 2001). Alternatively in 2004 Germany introduced the sustainability factor, whereas Japan implemented the demographic factors in the same year. Both factors operate more or less as an automatic balance mechanism. The automatic balance mechanism is to avoid any political difficulties by flexibly adapting the pension system to changing and unpredictable world.

We can introduce a guarantee pension (GP) to add on the NDC pensions. This is to provide an adequate income in old age. It should be financed by other sources than the contributions (payroll tax), since the policy objectives are quite different between the guarantee pension and the NDC one (see Figure 3.6).

3.8 Concluding remarks

The Japanese are increasingly concerned with the 'taste of pie' rather than the 'size of pie' or the 'distribution of pie.' When it comes to social security pensions, the most important question is whether or not they are worth buying. It has become a secondary concern how big or how fair they are. The basic design of the pension programme sshould be incentive-compatible. Contributions should be much more directly linked with old-age pension benefits, while elements of social adequacy should be incorporated in a separate tier of pension benefits financed by sources other than contributions.

Japan faces a dilemma that too many targets are sought to be achieved through a virtually single policy instrument of pensions. This contradicts the standard theory of policy assignment, which suggests that each policy objective can be best attained only if it is matched with each different policy instrument of comparative advantage. A diversified multi-tier system is thus most preferable.

Important, as well, is the separation of the legacy pension problem from the problem of re-building a sustainable and intergenerationally equitable pension system.

No one can claim to clearly see all the changes that lie in the decades ahead. Still the challenge is hard to ignore. Missing is a more explicit consideration of an automatic balancing mechanism for remedying possible mistakes in the projections toward the future.

Notes

1. A detailed explanation of the Japanese social security pension system is given by Takayama (1998, 2003b).
2. 1,000 yen = US$9.29 = euro 7.37 = UK £5.07 as at 18 May 2005.
3. A semi-annual bonus equivalent to 3.6 months salary is typically assumed.
4. The normal pensionable age of the KNH is 65, though Japan has special arrangements for a transition period between 2000 and 2025. See Takayama (2003b) for more details.
5. Excess liabilities of all social security pension programmes in Japan as at the end of March 2005 amounted to around 650 trillion yen, which is equivalent to 1.3 times the fiscal year 2004 GDP of Japan.
6. The amount of excess liabilities (EL) will vary depending on alternative discount rates. For example, a 2.1 per cent discount rate induces EL of 650 trillion yen, while another 4.0 per cent discount rate produces EL of 420 trillion yen. Part One excess liabilities can be termed as 'accrued-to-date net liabilities' or 'net termination liabilities'. See Franco (1995) and Holzmann et al. (2004).
7. Contributions to social security pensions operate as 'penalties on employment.' Further hikes in the contribution rate will seriously damage domestic companies which have been facing mega-competition on a global scale, thereby exerting negative effects on the economy, inducing a higher unemployment rate, lower economic growth, lower saving rates and so on. Further increases in the contribution rate will be sure to decrease take-home pay of actively working people in real terms, producing lower consumption and lower effective demand.
8. This section heavily depends on Takayama (2004).
9. Later retirement would be preferable for the country to achieve an active aging, if it has little substitution effects on employment for young people.
10. The excess liabilities partly arise from windfall gains given to the first generation in a pay-as-you-go pension system. This part should not be simply interpreted as 'the mistakes made in the past.'
11. Richard Musgrave once examined the credibility and long-run political viability of alternative contracts between generations, demonstrating that a '*Fixed Relative Position (FRP)*' approach will be most preferable (Musgrave, 1981). Following his suggestion, Germany and Japan had introduced a net indexation method in adjusting their social security pension benefits since the early 1990s. The FRP approach will face some difficulties, however. Among others, its approach could be only acceptable if the participation in the social security pension system will pay for the younger generations.

12. The payroll tax and the consumption-based tax might be indifferent in a steady-state economy, although they will induce different economic impacts in a transition period.
13. A 2 per cent earmarked consumption tax could be right, since the remaining excess liabilities of 90 trillion yen might be acceptable as a 'hidden' national debt.
14. Even if all the alternative measures above stated are implemented, currently young and future generations will still have to pay a substantial part of Part One excess liabilities. However, the current pensioners and baby-boomers should still try to do as much as possible to diminish the excess liabilities before any further increases in the contribution rate are considered.
15. A funded plan can be another alternative. However, it cannot escape from the so-called 'double burden' problem in the transition period, while the NDC is free from it.

References

Franco, D. (1995) 'Pension Liabilities: Their Use and Misuse in the Assessment of Fiscal Policies', European Commission, *Economic Papers*, No. 110.

Holzmann, R. et al. (2004) 'Implicit Pension Debt: Issues, Measurement and Scope in International Perspective', DP-0403, Social Protection Unit, World Bank.

Könberg, B. (2002) 'The Swedish Pension Reform: Some Lessons', DP-46, Project on Intergenerational Equity, Institute of Economic Research, Hitotsubashi University.

Musgrave, R.A. (1981) 'A Reappraisal of Financing Social Security', in Skidmore, F. (ed.), *Social Security Financing* (Cambridge: MIT Press), pp. 89–125.

Palmer, E. (2003) 'Pension Reform in Sweden', in Takayama (2003a, Chapter 7).

Project on Intergenerational Equity, Hitotsubashi University (2005) *The Balance Sheet of Social Security Pensions*, Proceedings, No. 6, PIE, Hitotsubashi University, Tokyo.

Settergren, O. (2001) 'The Automatic Balance Mechanism of the Swedish Pension System'. Available at www.rfv.se.

Takayama, N. (1998) *The Morning After in Japan: Its Declining Population, Too Generous Pensions and a Weakened Economy* (Tokyo: Maruzen).

Takayama, N. ed. (2003a) *Taste of Pie: Searching for Better Pension Provisions in Developed Countries* (Tokyo: Maruzen).

Takayama, N. (2003b) 'Pension Arrangements in the Oldest Country: The Japanese Case', in Takayama (2003a, Chapter 5).

Takayama, N. (2004) 'Changes in the Pension System', *Japan Echo*, vol. 31(5), October.

Part II
Ranking Infinite Utility Streams

4

A New Equity Condition for Infinite Utility Streams and the Possibility of being Paretian*

Geir B. Asheim, Tapan Mitra and Bertil Tungodden

4.1 Introduction

This chapter investigates the properties of a new equity condition for infinite utility streams. The condition, which was introduced in Asheim and Tungodden (2004a), is *Hammond Equity for the Future* (henceforth referred to as **HEF**), and it captures the following ethical intuition: A sacrifice by the present generation leading to a uniform gain for all future generations cannot lead to a less desirable utility stream if the present remains better off than the future even after the sacrifice.

In the terminology of Suzumura and Shinotsuka (2003), this new equity condition is a *consequentialist* condition, in the sense that it expresses preference for a more egalitarian distribution of utilities among generations. In contrast, the 'Weak Anonymity' condition, which often has been invoked to ensure equal treatment of generations (by requiring that any finite permutation of utilities should not change the social evaluation of the stream), is a purely *procedural* equity condition. As we discuss in Asheim and Tungodden (2004a), however, **HEF** is a very weak consequentialist condition. Under certain consistency requirements on the social preferences, it is not only weaker than the ordinary 'Hammond Equity' condition, but it is also implied by other consequentialist equity conditions like the Pigou–Dalton principle of transfers and the Lorenz Domination principle.

From Koopmans (1960), Diamond (1965), and later contributions (e.g., Svensson, 1980; Shinotsuka, 1998; Basu and Mitra, 2003; Fleurbaey

* We thank Aanund Hylland and the participants in the 2005 IEA Roundtable Meeting on Intergenerational Equity in Hakone for discussions and comments. Asheim gratefully acknowledges the hospitality of Cornell University and financial support received from the Research Council of Norway.

and Michel, 2003; Sakai, 2003; Xu, 2005) we know that it is problematic in the context of infinite utility streams to combine procedural *equity* conditions with conditions ensuring the *efficiency* of a socially preferred utility stream. In particular, Diamond (1965) states the result that the 'Weak Anonymity' condition cannot be combined with the 'Strong Pareto' condition when the social preferences are complete, transitive and continuous in the sup norm topology (a result that he attributes to M.E. Yaari). This impossibility result has subsequently been strengthened in several ways. The inconsistency remains even if 'Strong Pareto' is replaced by 'Weak Pareto' (Fleurbaey and Michel, 2003) or 'Sensitivity To the Present' (Sakai, 2003). Moreover, Diamond's (1965) proof does not use the full force of the assumption that the social preferences are complete, transitive and continuous in the sup norm topology, and Basu and Mitra (2003a) show that the inconsistency remains even if this assumption is replaced by an assumption of numerical representability.

Suzumura and Shinotsuka (2003) and Sakai (2006) show that the same kind of impossibility results can be established when consequentialist equity conditions are combined with 'Strong Pareto'. In particular, Suzumura and Shinotsuka (2003) establish that the Lorenz Domination principle is not compatible with 'Strong Pareto' when social preferences are upper semi-continuous in the sup norm topology.

The investigations by Suzumura and Shinotsuka (2003) and Sakai (2006) encourage us to carry out a similar analysis for our condition **HEF**. Since **HEF** is a weak condition when compared to other consequentialist equity conditions, it is of interest to establish whether it to a greater extent can be combined with Paretian conditions. We show in this chapter that, unfortunately, this is not the case: Condition **HEF** is not compatible with 'Strong Pareto' when social preferences are upper semi-continuous in the sup norm topology. Both our result and the corresponding result by Suzumura and Shinotsuka (2003) do not require any consistency requirements (like completeness and transitivity) on the social preferences. However, if we impose that the social preferences are complete, transitive and continuous in the sup norm topology, and satisfy an 'Independent Future' condition, then **HEF** cannot even be combined with the 'Weak Pareto' condition. These are discouraging results, given the weakness of **HEF** and its possible ethical appeal.

Our chapter is organized as follows. In section 4.2 we present the setting, and state the conditions that we return to in later sections. In section 4.3 we show under what circumstances **HEF** is implied by other consequentialist equity conditions. In section 4.4 we establish a basic impossibility result, on which the findings in the subsequent sections will be based. In section 4.5 we show that **HEF** cannot be combined with 'Strong Pareto' when preferences satisfy a restricted form for upper semi-continuity in the sup norm topology, while in section 4.6 we report on the inconsistency with 'Weak Pareto' and 'Sensitivity To the Present' under additional conditions. Results relating to the Pigou-Dalton and Lorenz Domination principles are reported as corollaries. In section 4.7 we present examples that serve to clarify the role of the various conditions in the impossibility results that arise in the present framework. Finally, in section 4.8 we discuss what these negative results entail for the usefulness of condition **HEF** and other consequentialist equity conditions as ethical guidelines for intergenerational equity.

4.2 Framework and conditions

Let $\Re$ be the set of real numbers and $\aleph$ the set of positive integers. The set of infinite utility streams is $X = Y^{\aleph}$, where $[0, 1] \subseteq Y \subseteq \Re$. Denote by ${}_1\mathbf{u} = (u_1, u_2, \ldots, u_t, \ldots)$ an element of X, where u_t is the utility of generation t, and denote by ${}_1\mathbf{u}_T = (u_1, u_2, \ldots, u_T)$ and ${}_{T+1}\mathbf{u} = (u_{T+1}, u_{T+2}, \ldots)$ the T-head and T-tail of the utility stream respectively. Write ${}_{\text{con}}w = (w, w, \ldots)$ for a stream with a constant level of utility equal to $w \in Y$. Throughout this chapter we assume at least ordinally measurable level comparable utilities; i.e., what Blackorby, Donaldson and Weymark (1984) refer to as 'level-plus comparability'.

For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, we write ${}_1\mathbf{u} \geq {}_1\mathbf{v}$ if and only if $u_t \geq v_t$ for all $t \in \aleph$; ${}_1\mathbf{u} > {}_1\mathbf{v}$ if and only if ${}_1\mathbf{u} \geq {}_1\mathbf{v}$ and ${}_1\mathbf{u} \neq {}_1\mathbf{v}$; and ${}_1\mathbf{u} >> {}_1\mathbf{v}$ if and only if $u_t > v_t$ for all $t \in \aleph$.

Social preferences are a binary relation R on X, where for any ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, ${}_1\mathbf{u}\ R\ {}_1\mathbf{v}$ entails that ${}_1\mathbf{u}$ is deemed socially at least as good as ${}_1\mathbf{v}$. Denote by I and P the symmetric and asymmetric parts of R; i.e., ${}_1\mathbf{u}\ I\ {}_1\mathbf{v}$ is equivalent to ${}_1\mathbf{u}\ R\ {}_1\mathbf{v}$ and ${}_1\mathbf{v}\ R\ {}_1\mathbf{u}$ and entails that ${}_1\mathbf{u}$ is deemed socially indifferent to ${}_1\mathbf{v}$, while ${}_1\mathbf{u} P {}_1\mathbf{v}$ is equivalent to ${}_1\mathbf{u}\ R\ {}_1\mathbf{v}$ and $\neg {}_1\mathbf{v}\ R\ {}_1\mathbf{u}$ and entails that ${}_1\mathbf{u}$ is deemed socially preferable to ${}_1\mathbf{v}$. We will consider different sets of conditions on R.

We consider two *consistency* conditions.

Condition O (*Order*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, ${}_1\mathbf{u}\ R\ {}_1\mathbf{v}$ or ${}_1\mathbf{v}\ R\ {}_1\mathbf{u}$. For all ${}_1\mathbf{u}$, ${}_1\mathbf{v}$, ${}_1\mathbf{w} \in X$, ${}_1\mathbf{u}\ R\ {}_1\mathbf{v}$ and ${}_1\mathbf{v}\ R\ {}_1\mathbf{w}$ imply ${}_1\mathbf{u}\ R\ {}_1\mathbf{w}$.

Condition QT (*Quasi-Transitivity*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v}$, ${}_1\mathbf{w} \in X$, ${}_1\mathbf{u}\ P\ {}_1\mathbf{v}$ and ${}_1\mathbf{v}\ P\ {}_1\mathbf{w}$ imply ${}_1\mathbf{u}\ P\ {}_1\mathbf{w}$.

Condition **O** implies condition **QT**, while the converse does not hold.

We consider four *continuity* conditions (relative to the sup norm topology). For the results of the present chapter, it is sufficient to use the restricted forms, where we only use the sup norm to compare with streams that are eventually constant. Such restricted continuity conditions are less demanding than their non-restricted counterparts.

Condition C (*Continuity*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if $\lim_{n\to\infty} \sup_t |u_t^n - u_t| = 0$ with, for all n, $\neg {}_1\mathbf{v}\ P\ {}_1\mathbf{u}^n$ (resp. $\neg {}_1\mathbf{u}^n\ P\ {}_1\mathbf{v}$), then $\neg {}_1\mathbf{v}\ P\ {}_1\mathbf{u}$ (resp. $\neg {}_1\mathbf{u}\ P\ {}_1\mathbf{v}$).

Condition RC (*Restricted Continuity*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if there exists $T \geq 1$ such that $u_t = w$ for all $t \geq T$, and $\lim_{n\to\infty} \sup_t |u_t^n - u_t| = 0$ with, for all n, $\neg {}_1\mathbf{v}\ P\ {}_1\mathbf{u}^n$ (resp. $\neg {}_1\mathbf{u}^n P\ {}_1\mathbf{v}$), then $\neg {}_1\mathbf{v}\ P\ {}_1\mathbf{u}$ (resp. $\neg {}_1\mathbf{u}\ P\ {}_1\mathbf{v}$).

Condition USC (*Upper Semi-Continuity*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if $\lim_{n\to\infty} \sup_t |u_t^n - u_t| = 0$ with, for all n, $\neg {}_1\mathbf{v}\ P\ {}_1\mathbf{u}^n$, then $\neg {}_1\mathbf{v}\ P\ {}_1\mathbf{u}$.

Condition RUSC (*Restricted Upper Semi-Continuity*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if there exists $T \geq 1$ such that $u_t = w$ for all $t \geq T$, and $\lim_{n\to\infty} \sup_t |u_t^n - u_t| = 0$ with, for all n, $\neg {}_1\mathbf{v}\ P\ {}_1\mathbf{u}^n$, then $\neg {}_1\mathbf{v}\ P\ {}_1\mathbf{u}$.

Condition **C** implies conditions **RC** and **USC**, while each of the latter implies condition **RUSC**. The converses do not hold.

We consider eight *efficiency* conditions. The first four are *Paretian* conditions, where condition **WD** has been analyzed by Basu and Mitra (2007b), while condition **RWP** is used by Asheim and Tungodden (2004a) (but referred to there as 'Sensitivity').

Condition SP (*Strong Pareto*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if ${}_1\mathbf{u} > {}_1\mathbf{v}$, then ${}_1\mathbf{u}\ P\ {}_1\mathbf{v}$.

Condition WD (*Weak Dominance*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if there exists $s \geq 1$ such that $u_s > v_s$ and $u_t = v_t$ for $t \neq s$, then ${}_1\mathbf{u}\ P\ {}_1\mathbf{v}$.

Condition WP (*Weak Pareto*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if ${}_1\mathbf{u} >> {}_1\mathbf{v}$, then ${}_1\mathbf{u}\ P\ {}_1\mathbf{v}$.

Condition RWP (*Restricted Weak Pareto*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if ${}_1\mathbf{u} >> {}_1\mathbf{v}$ and there exists $T \geq 1$ such that $u_t = w$ and $v_t = x$ for all $t \geq T$, then ${}_1\mathbf{u}\ P\ {}_1\mathbf{v}$.

Condition **SP** implies conditions **WD** and **WP**, while the converses do not hold. Moreover, condition **WP** implies condition **RWP**, while the converse does not hold.

The remaining four are *sensitivity* conditions, where condition **STP** has been analyzed by Sakai (2003), while condition **WS** coincides with Koopmans' (1960) postulate 2.

Condition SS (*Strong Sensitivity*) For all ${}_2\mathbf{w} \in X$, there exist u_1, $v_1 \in Y$ with $u_1 > v_1$ such that $(u_1, {}_2\mathbf{w})\ P (v_1, {}_2\mathbf{w})$.

Condition STP (*Sensitivity To the Present*) For all ${}_1\mathbf{w} \in X$, there exist ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, and $T \geq 1$ such that $({}_1\mathbf{u}_T, {}_{T+1}\mathbf{w})\ P({}_1\mathbf{v}_T, {}_{T+1}\mathbf{w})$.

Condition RS (*Restricted Sensitivity*) There exist u, $v \in Y$ with $u > v$ such that $(u, {}_{\text{con}}v) P (v, {}_{\text{con}}v)$.

Condition WS (*Weak Sensitivity*) There exist u_1, $v_1 \in Y$ and ${}_2\mathbf{w} \in X$ such that $(u_1, {}_2\mathbf{w})\ P(v_1, {}_2\mathbf{w})$.

Condition **WD** implies condition **SS**, which in turn implies conditions **STP** and **RS**, while the converses do not hold. Furthermore, condition **RS** implies condition **WS**, while the converse does not hold.

Finally, we consider four consequentialist *equity* conditions. The two first require only, as we assume throughout this chapter, at least ordinally measurable level comparable utilities. For complete social preferences these conditions coincide with those suggested by Hammond (1976) and Asheim and Tungodden (2004a), respectively.

Condition HE (*Hammond Equity*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if ${}_1\mathbf{u}$ and ${}_1\mathbf{v}$ satisfy that there exists a pair r and s such that $u_r > v_r > v_s > u_s$ and $v_t = u_t$ for $t \neq r, s$, then $\neg {}_1\mathbf{u}\ P\ {}_1\mathbf{v}$.

Condition HEF (*Hammond Equity for the Future*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if ${}_1\mathbf{u}$ and ${}_1\mathbf{v}$ satisfy that $u_1 > v_1 > \mathrm{x} > w$ and $u_t = w$ and $v_t = \mathrm{x}$ for all $t > 1$, then $\neg {}_1\mathbf{u}\ P\ {}_1\mathbf{v}$.

The two next equity conditions require, in addition, that utilities are at least cardinally measurable and unit comparable. Such consequentialist equity conditions have been used in the context of infinite streams by, e.g., Birchenhall and Grout (1979), Asheim (1991), and Fleurbaey and Michel (2001), as well as Suzumura and Shinotsuka (2003) and Sakai (2006). The former of the two conditions below is in the exact form suggested by Suzumura and Shinotsuka (2003).

Condition WLD (*Weak Lorenz Domination*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if ${}_1\mathbf{u}$ and ${}_1\mathbf{v}$ satisfy that there exist $T > 1$ such that ${}_1\mathbf{v}_{\mathrm{T}}$ Lorenz dominates ${}_1\mathbf{u}_{\mathrm{T}}$ and ${}_{T+1}\mathbf{u} = {}_{T+1}\mathbf{v}$, then $\neg {}_1\mathbf{u}\ P\ {}_1\mathbf{v}$.

Condition WPD (*Weak Pigou-Dalton*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if ${}_1\mathbf{u}$ and ${}_1\mathbf{v}$ satisfy that there exist a positive number ε and a pair r and s such that $u_r - \varepsilon = v_r \geq v_s = u_s + \varepsilon$ and $v_t = u_t$ for $t \neq r, s$, then $\neg {}_1\mathbf{u}\ P\ {}_1\mathbf{v}$.

Condition **WLD** implies condition **WPD**, while the converse does not hold. The implications between condition **HEF**, on the one hand, and the three other equity conditions, on the other hand, are treated in the next section. All results in this chapter would still hold if we replaced condition **WPD** by a weaker rank-preserving version where the premise requires also that, for $t \neq r, s$, $v_r \geq v_t$ if $u_r \geq u_t$ and $v_t \geq v_s$ if $u_t \geq u_s$ (cf. Fields and Fei, 1978).

We end this section by stating a condition which is implied by Koopmans' (1960) postulates 3b and 4. It means that a decision concerning only generations from the second period on can be made as if the present time (period 1) was actually at period 2; i.e., as if generations $\{1, 2, \dots\}$ would have taken the place of generations $\{2, 3, \dots\}$. It is stated by this name, but in a slightly stronger form, by Fleurbaey and Michel (2003).

Condition IF (*Independent Future*) For all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$ with $u_1 = v_1$, ${}_1\mathbf{u}\ R_1\ \mathbf{v}$ if and only if ${}_2\mathbf{u}\ R_2\ \mathbf{v}$.

4.3 Hammond *Equity for the Future*

For streams where utility is constant from the second period on, condition **HEF** states the following: If the present is better off than the future and a sacrifice now leads to a uniform gain for all future generations, then such a transfer from the present to the future cannot lead to a stream that is less desirable in social evaluation, as long as the present remains better off than the future.

To appreciate the weakness of condition **HEF**, consider the following result.

Proposition 1 *Let* $Y \supseteq [0, 1]$. *If* ***QT*** *and* ***RWP*** *hold, then each of* ***HE*** *and* ***WLD*** *implies* ***HEF***. *If* ***O*** *and* ***RWP*** *hold, then* ***WPD*** *implies* ***HEF***.

Proof. Assume $u'' > u' > w' > w''$. We must show under the given conditions that each of **HE**, **WLD**, and **WPD** implies $\neg(u'',_{\text{con}} w'')\ P\ (u', {}_{\text{con}}w')$.

Since $u'' > u' > w' > w''$, there exists an integer $T \geq 1$ and utilities $v, x \in Y$ satisfying $u'' > u' > v \geq w' > x > w''$ and $u'' - v = T(x - w'')$.

If **HE** holds, then $\neg(u'', {}_{\text{con}}w'')\ P\ (v, x, {}_{\text{con}}w'')$, and by **RWP**, $(u', {}_{\text{con}}w')\ P\ (v, x, {}_{\text{con}}w'')$. By **QT**, $\neg(u'', {}_{\text{con}}w'')\ P\ (u', {}_{\text{con}}w')$.

Consider next **WLD** and **WPD**. Let ${}_1\mathbf{u}^0 = (u'', {}_{\text{con}}w'')$, and define, for $n \in \{1, \ldots, T\}$, ${}_1\mathbf{u}^n$ inductively as follows:

$$
\begin{aligned}
u_t^n &= u_t^{n-1} - (x - w'') && \text{for } t = 1 \\
u_t^n &= x && \text{for } t = 1 + n \\
u_t^n &= u_t^{n-1} && \text{for } t \neq 1, 1 + n.
\end{aligned}
$$

If **WLD** holds, then $\neg{}_1\mathbf{u}^0\ P\ {}_1\mathbf{u}^T$, and by **RWP**, $(u', {}_{\text{con}}w')\ P\ {}_1\mathbf{u}^T$. By **QT**, $\neg(u'', {}_{\text{con}}w'')\ P\ (u', {}_{\text{con}}w')$ since ${}_1\mathbf{u}^0 = (u'', {}_{\text{con}}w'')$.

If **WPD** holds, then by **O**, for $n \in \{1, \ldots, T\}$, ${}_1\mathbf{u}^n\ R\ {}_1\mathbf{u}^{n-1}$, and by **RWP**, $(u', {}_{\text{con}}w')\ P\ {}_1\mathbf{u}^T$. By **O**, $(u', {}_{\text{con}}w')\ P\ (u'', {}_{\text{con}}w'')$ since ${}_1\mathbf{u}^0 = (u'', {}_{\text{con}}w'')$. Hence, $\neg(u'', {}_{\text{con}}w'')\ P\ (u', {}_{\text{con}}w')$. □

Note that condition **HEF** involves a comparison between a sacrifice by a single generation and a uniform gain for each member of an infinite set of generations that are worse off. Hence, contrary to the standard 'Hammond Equity' condition, if utilities are made (at least) cardinally measurable and fully comparable, then the transfer from the better-off present to the worse-off future specified in condition **HEF** increases the sum of utilities obtained by summing the utilities of a sufficiently large number T of generations. This entails that condition **HEF** is implied by both the Pigou–Dalton principle of transfers and the Lorenz Domination principle, independently of what specific cardinal utility scale is imposed (provided that the consistency conditions specified in Proposition 1 are satisfied). Hence, 'Hammond Equity for the Future' can be endorsed both from an egalitarian and utilitarian point of view. In particular, condition **HEF** is much weaker and more compelling than the standard 'Hammond Equity' condition.

4.4 Basic impossibility result

In the present section we establish that **HEF** is in direct conflict with **RS** under **RUSC**. Hence, there are no restricted sensitive and restricted upper semi-continuous social preferences that satisfy our new equity condition. In the subsequent two sections we note how **RS** is implied by various efficiency conditions. Proposition 2 is thereby used to show how **HEF** cannot be combined with efficiency conditions as long as specific forms of continuity are imposed.

Proposition 2 *Let* $Y \supseteq [0, 1]$. *There are no social preferences satisfying* ***RUSC****,* ***RS****, and* ***HEF****.*

Proof. Suppose there exist social preferences R satisfying **RUSC**, **RS**, and **HEF**.

Step 1: By **RS**, there exists $u, v \in Y$ with $u > v$ such that $(u, {}_{\text{con}}v)\ P\ (v, {}_{\text{con}}v)$. Define $a = u - v$. We claim that there is $b \in (0, a)$ such that

$$(u, {}_{\text{con}}v) P (v + b, {}_{\text{con}}v).$$

If not, for every $b \in (0, a)$ we have $\neg(u, {}_{\text{con}}v)\ P\ (v + b, {}_{\text{con}}v)$. By letting $b \to 0$, we have by **RUSC**: $\neg(u, {}_{\text{con}}v)\ P\ (v, {}_{\text{con}}v)$. This contradicts $(u, {}_{\text{con}}v)\ P\ (v, {}_{\text{con}}v)$ and establishes our claim.

Step 2: For every $c \in (0, b)$, noting that $u > v + b > v + c > v$, **HEF** implies that $\neg(u, {}_{\text{con}}v)\ P\ (v + b, {}_{\text{con}}(v + c))$. By letting $c \to 0$ and using **RUSC**, we get

$$\neg(u, {}_{\text{con}}v) P (v + b, {}_{\text{con}}v).$$

This contradicts the claim proved in Step 1, and establishes the proposition. □

Note that no consistency conditions (like completeness and transitivity) on the social preferences are required for this result

The Diamond–Yaari impossibility result (Diamond, 1965) states that conditions **C** and **SP** are inconsistent with 'Weak Anonymity' under the additional assumptions of completeness and transitivity. Actually, the proof provided allows **C** to be replaced by lower semi-continuity and **SP** to be replaced by **WD**, and with a different proof than the one given by Diamond (1965) one can even replace lower semi-continuity by **USC**. Compared to this result, we claim that it is equally worrying that the even weaker conditions **RUSC** and **RS** are inconsistent with assigning priority to an infinite number of worst-off generations in comparisons where the assignment of such priority only reduces the utility of the better-off present generation, as expressed by condition **HEF**. In this respect, note that **HEF** neither implies nor is implied by 'Weak Anonymity', and thus Proposition 2 is different from impossibility results based on 'Weak Anonymity' as a procedural equity condition.

4.5 Strong Pareto

Since **SP** implies **RS**, it is a straightforward implication of Proposition 2 that **HEF** is in direct conflict with **SP** under **RUSC**. Hence, there are no strongly Paretian and restricted upper semi-continuous social preferences that satisfy 'Hammond Equity for the Future'.

Proposition 3 *Let $Y \supseteq [0, 1]$. There are no social preferences satisfying **RUSC**, **SP**, and **HEF**.*

Since **SP** implies **RWP**, we obtain the following corollary by combining Propositions 1 and 3.

Corollary 1 *Let $Y \supseteq [0, 1]$. If **QT** holds, then there are no social preferences satisfying **RUSC**, **SP**, and **HE**; or **RUSC**, **SP**, and **WLD**. If **O** holds, then there are no social preferences satisfying **RUSC**, **SP**, and **WPD**.*

It should be remarked that the results of Corollary 1 are available in other variants; in particular, it follows from Theorem 3 of Suzumura and Shinotsuka (2003) that condition **QT** is not needed for showing that there are no social preferences satisfying **USC**, **SP**, and **WLD**. Moreover, both Suzumura and Shinotsuka (2003, Theorem 1) and Sakai (2006, Theorem 2) show that only condition **QT** is needed for **USC** and **SP** to be incompatible with a strengthened version of **WPD** (namely, for all ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, if ${}_1\mathbf{u}$ and ${}_1\mathbf{v}$ satisfy that there exist a positive number ε and a pair r and s such that $u_r - \varepsilon = v_r \geq v_s = u_s + \varepsilon$ and $v_t = u_t$ for $t \neq r$, s, then ${}_1\mathbf{v}\ P_1\mathbf{u}$).

4.6 Weaker Paretian conditions

We now show that **HEF** is even in conflict with **WP**, provided that the social preferences satisfy conditions **O**, **RC**, and **IF**. Hence, there are no weakly Paretian, complete, transitive and restricted continuous social preferences that satisfy both 'Independent future' and our new equity condition.

Proposition 4 *Let $Y \supseteq [0, 1]$. If **O** and **IF** hold, then there are no social preferences satisfying **RC**, **WP**, and **HEF**.*

Proposition 4 follows by combining Proposition 2 with the following lemma.

Lemma 1 *Let $Y \supseteq [0, 1]$, and assume that the social preferences R satisfy **O**, **RC**, **WP**, and **IF**. Then the social preferences R satisfy **RS**.*

Proof. Assume that the social preferences R satisfy **O**, **RC**, **WP**, and **IF**. Consider the stream ${}_1\mathbf{u} \in X$ defined by, for all $t \geq 1$, $u_t = 1/t$; i.e.

$$ {}_1\boldsymbol{u} = \left(1, \tfrac{1}{2}, \tfrac{1}{3}, \tfrac{1}{4}, \ldots\right). $$

By **WP**, ${}_{\text{con}}1\ P\ {}_2\mathbf{u}\ P\ {}_{\text{con}}0$. By **O** and **WP**, there exists $w \in [0,\ 1]$ such that $w = \inf\{x|_{\text{con}}x\ R\ {}_2\mathbf{u}\} = \sup\{x|_2\mathbf{u}\ R\ {}_{\text{con}}x\}$. By **O** and **RC**, ${}_{\text{con}}w\ I\ {}_2\mathbf{u}$ and $w \in (0,\ 1)$. By **IF**, $(1,\ {}_{\text{con}}w)\ I\ {}_1\mathbf{u}$. Since, by **WP**, ${}_1\mathbf{u}\ P\ {}_2\mathbf{u}$, we have that $(1,\ {}_{\text{con}}w)\ I\ {}_1\mathbf{u}\ P\ {}_2\mathbf{u}\ I\ (w,\ {}_{\text{con}}w)$. Hence, by **O**, $(1,\ {}_{\text{con}}w)\ P\ (w,\ {}_{\text{con}}w)$, where $1 > w$. This shows that R satisfies **RS**. □

Since **O** implies **QT** and **WP** implies **RWP**, we obtain the following corollary by combining Propositions 1 and 4.

Corollary 2 *Let* $Y \supseteq [0,\ 1]$. *If **O** and **IF** hold, then there are no social preferences satisfying **RC**, **WP**, and **HE**; or **RC**, **WP**, and **WLD**; or **RC**, **WP**, and **WPD**.*

Moreover, as the following proposition establishes, **HEF** is also in conflict with **STP** and **RWP**, provided that the social preferences satisfy conditions **O**, **RC**, and **IF**.

Proposition 5 *Let* $Y = [0,\ 1]$. *If **O** and **IF** hold, then there are no social preferences satisfying **RC**, **STP**, **RWP**, and **HEF**.*

Proposition 5 follows by combining Proposition 2 with Lemma 3 below. The proof of Lemma 3 makes use of the following result.

Lemma 2 *Let* $Y = [0,\ 1]$, *and assume that the social preferences R satisfy **O**, **RUSC**, and **RWP**. Then, for all* ${}_1\mathbf{u} \in X$ *and all* $T \geq 1$, $({}_1\mathbf{u}_T,\ {}_{\text{con}}0)\ R\ {}_{\text{con}}0$.

Proof. Assume that the social preferences R satisfy **O**, **RUSC**, and **RWP**. Let ${}_1\mathbf{u} \in X$. For $a \in (0,\ 1)$, define ${}_1\mathbf{u}(a)$ as follows: $u_t(a) = u_t + a(1 - u_t)$ for $t = 1, \ldots, T$, and $u_t(a) = a$ for $t > T$. For each $a \in (0,\ 1)$, ${}_1\mathbf{u}(a) \in X$, with $u_t(a) \geq a > 0$ for $t = 1, \ldots, T$, and $u_t(a) = a > 0$ for $t > T$. By **RWP**, ${}_1\mathbf{u}(a)\ P\ {}_{\text{con}}0$ for each $a \in (0,\ 1)$. Letting $a \to 0$ and using **O** and **RUSC**, we get $({}_1\mathbf{u}_T,\ {}_{\text{con}}0)$ $R_{\text{con}}0$. □

Lemma 3 *Let* $Y = [0,\ 1]$, *and assume that the social preferences R satisfy **O**, **RC**, **STP**, **RWP**, and **IF**. Then the social preferences R satisfy **RS**.*

Proof. Suppose that the social preferences R satisfy **O**, **RC**, **STP**, **RWP**, and **IF**, but violate **RS**. Since **RS** is violated, we must have

$$\neg(1,\ {}_{\text{con}}0)P(0,\ {}_{\text{con}}0).$$

Step 1: By **O**, we have $(0,\ {}_{\text{con}}0)\ R\ (1,\ {}_{\text{con}}0)$. On the other hand, by Lemma 2, $(1,\ {}_{\text{con}}0)\ R\ (0,\ {}_{\text{con}}0)$, since **O** and **RC** (and thus, **RUSC**) hold. Hence, we must have

$$(1,\ {}_{\text{con}}0)I(0,\ {}_{\text{con}}0).$$

Define

$$
\begin{aligned}
{}_1\mathbf{x}^0 &= (0, 0, 0, 0, \dots) \\
{}_1\mathbf{x}^1 &= (1, 0, 0, 0, \dots) \\
{}_1\mathbf{x}^2 &= (1, 1, 0, 0, \dots) \\
{}_1\mathbf{x}^3 &= (1, 1, 1, 0, \dots)
\end{aligned}
$$

and so forth. We have already established that ${}_1\mathbf{x}^1 \; I \; {}_1\mathbf{x}^0$. Furthermore, by **IF**, for all $n \in \aleph$, ${}_1\mathbf{x}^n \; I \; {}_1\mathbf{x}^{n-1}$ implies ${}_1\mathbf{x}^{n+1} \; I \; {}_1\mathbf{x}^n$. Since **O** holds, it follows by induction that, for all $n \in \aleph$, ${}_1\mathbf{x}^n \; I \; (0, {}_{\text{con}}0)$.

Step 2: Using **STP**, there exist ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$, and $T \geq 1$ such that

$$(u_1, \dots, u_T, {}_{\text{con}}0) \, P \, (v_1, \dots, v_T, {}_{\text{con}}0).$$

By **O** and **RC**, there exists $b \in (0, 1)$ such that

$$(bu_1, \dots, bu_T, {}_{\text{con}}0) \, P \, (v_1, \dots, v_T, {}_{\text{con}}0).$$

For $c \in (0, 1)$, define ${}_1\mathbf{w}(c)$ as follows: $w_t(c) = 1$ for $t = 1, \dots, T$, and $w_t(c) = c$ for $t > T$. Then, by **RWP**, we have ${}_1\mathbf{w}(c) \; P \; (bu_1, \dots, bu_T, {}_{\text{con}}0)$ for each $c \in (0, 1)$. Letting $c \to 0$, and using **O** and **RC**, we have

$${}_1\mathbf{x}^T R (bu_1, \dots, bu_T, {}_{\text{con}}0).$$

On the other hand, by Lemma 2,

$$(v_1, \dots, v_T, {}_{\text{con}}0) \, R \, (0, {}_{\text{con}}0),$$

since **O** and **RC** (and thus, **RUSC**) hold. Hence, ${}_1\mathbf{x}^T \; R \; (bu_1, \dots, bu_T, {}_{\text{con}}0)$ $P (v_1, \dots, v_T, {}_{\text{con}}0) \; R \; (0, {}_{\text{con}}0)$, and using **O** we get

$${}_1\mathbf{x}^T \, P \, (0, {}_{\text{con}}0).$$

This contradicts the conclusion reached in Step 1, and establishes the proposition. □

Since **O** implies **QT**, we obtain the following corollary by combining Propositions 1 and 5.

Corollary 3 *Let $Y = [0, 1]$. If **O** and **IF** hold, then there are no social preferences satisfying **RC**, **STP**, **RWP**, and **HE**; or **RC**, **STP**, **RWP**, and **WLD**; or **RC**, **STP**, **RWP**, and **WPD**.*

4.7 Examples

We discuss three examples of social preferences, which clarify the role of the various conditions in the impossibility results arising in the framework of this chapter.

The first example provides an instance of social preferences which satisfy conditions **O**, **RC**, **WP** (and thus, **RWP**), **STP**, and **HEF**. This possibility result points to the critical role played by 'Restricted Sensitivity' (**RS**) in the impossibility result stated in Proposition 2, and the role played by 'Independent Future' (**IF**) in the impossibility results stated in Propositions 4 and 5. The example does not satisfy **RS**, and it does not satisfy **IF** (as is to be expected, since **IF**, in conjunction with the other conditions, would imply that **RS** hold, as we have shown in Lemmas 1 and 3 of this chapter).

The second example provides an instance of *representable* social preferences which satisfy both **RS** and **HEF**. This example does not satisfy **RUSC** since, by Proposition 2, **RS** and **HEF** imply that **RUSC** does not hold. The possibility result that Example 2 constitutes indicates that the even the very weak continuity condition **RUSC**, used in the impossibility result of Proposition 2, is a strong restriction.

The third example provides an instance of social preferences which satisfy conditions **O**, **RC**, **RWP**, **HEF**, and **IF**. This example satisfies neither **WP** nor **STP** since, by Propositions 4 and 5, the other conditions imply that **WP** and **STP** do not hold. Hence, this result illustrates the important role played by **WP** in Proposition 4 and **STP** in Proposition 5.

All three examples indicate that the notion of equity captured by **HEF** is a very weak one. It would be difficult to argue that the social preferences presented in these examples are, in any reasonable sense, 'equitable'. Hence, **HEF** is designed to be a *necessary* condition for equity, classifying as 'inequitable' social preferences that do *not* satisfy the condition.

Example 1 Let $Y \supseteq [0, 1]$, and define, for each ${}_1\mathbf{u} \in X$, $W({}_1\mathbf{u}) = u_2$. Now, define R by

$$\text{for all } {}_1\mathbf{u},\ {}_1\mathbf{v} \in X, {}_1\mathbf{u}\ R_1\mathbf{v} \text{ if and only if } W({}_1\mathbf{u}) \geq W({}_1\mathbf{v}).$$

Hence, the social preferences R are represented by the social welfare function W: $X \to Y$. Then the social preferences R satisfy **O**. They also satisfy **RC** and **WP**. To verify **HEF**, let ${}_1\mathbf{u}, {}_1\mathbf{v} \in X$ satisfy that $u_1 > v_1 > x > w$ and $u_t = w$ and $v_t = x$ for all $t > 1$. Then $W({}_1\mathbf{u}) = w$ and $W({}_1\mathbf{v}) = x$. Thus, $W({}_1\mathbf{v}) > W({}_1\mathbf{u})$, and so ${}_1\mathbf{v}\ P\ {}_1\mathbf{u}$. Finally, one can check that **STP** is satisfied as follows. Given any ${}_1\mathbf{w} \in X$, choose ${}_1\mathbf{u} = {}_{\text{con}}1$ and ${}_1\mathbf{v} = {}_{\text{con}}0$, and $T = 2$. Then we have $({}_1\mathbf{u}_T, {}_{T+1}\mathbf{w}) = (1, 1, {}_{T+1}\mathbf{w})$ and $({}_1\mathbf{v}_T, {}_{T+1}\mathbf{w}) = (0, 0, {}_{T+1}\mathbf{w})$. Thus, $W({}_1\mathbf{u}_T, {}_{T+1}\mathbf{w}) = 1$ and $W({}_1\mathbf{v}_T, {}_{T+1}\mathbf{w}) = 0$, so that $({}_1\mathbf{u}_T, {}_{T+1}\mathbf{w})\ P\ ({}_1\mathbf{v}_T, {}_{T+1}\mathbf{w})$. Clearly, R violates **RS** and **IF**.

Example 2 Let $Y = [0, 1]$, and define, for ${}_1\mathbf{u} \in X$, ${}_1\mathbf{u} \neq {}_{\text{con}}0$, $W({}_1\mathbf{u}) = 1$; and define $W({}_{\text{con}}0) = 0$. Now, define R by

$$\text{for all } {}_1\mathbf{u},\ {}_1\mathbf{v} \in X, {}_1\mathbf{u} R {}_1\mathbf{v} \text{ if and only if } W({}_1\mathbf{u}) \geq W({}_1\mathbf{v}).$$

Hence, the social preferences R are represented by the social welfare function $W: X \to \{0, 1\}$. Then the social preferences R satisfy **O**. To verify **HEF**, let ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$ satisfy that $u_1 > v_1 > x > w$ and $u_t = w$ and $v_t = x$ for all $t > 1$. Then $u_1 \neq 0$ so ${}_1\mathbf{u} \neq {}_{\text{con}}0$ and, consequently, $W({}_1\mathbf{u}) = 1$. Also $v_1 \neq 0$ so ${}_1\mathbf{v} \neq {}_{\text{con}}0$ and consequently, $W({}_1\mathbf{v}) = 1$. Then $W({}_1\mathbf{u}) = W({}_1\mathbf{v})$, and so ${}_1\mathbf{v}\ I\ {}_1\mathbf{u}$. To verify **RS**, choose $u = 1$ and $v = 0$. Then $(u, {}_{\text{con}}v) = (1, {}_{\text{con}}0)$ and $(v, {}_{\text{con}}v) = (0, {}_{\text{con}}0)$. Thus, $W(u, {}_{\text{con}}v) = 1$ and $W(v, {}_{\text{con}}v) = 0$ so that $(u, {}_{\text{con}}v)\ P\ (v, {}_{\text{con}}v)$. Clearly, R violates **RUSC**.

Example 3 Let $Y = [0, 1]$, and define, for each ${}_1\mathbf{u} \in X$,

$$W({}_1\mathbf{u}) = \lambda \limsup_{t\to\infty} u_t + (1 - \lambda) \liminf_{t\to\infty} u_t, \text{ where } 0 \leq \lambda \leq 1.$$

Now, define R by

$$\text{for all } {}_1\mathbf{u}, {}_1\mathbf{v} \in X, {}_1\mathbf{u} R {}_1\mathbf{v} \text{ if and only if } W({}_1\mathbf{u}) \geq W({}_1\mathbf{v}).$$

Hence, the social preferences R are represented by the social welfare function $W: X \to [0, 1]$. If $0 < \lambda < 1$, then the social preferences R presume that utilities are (at least) cardinally measurable and fully comparable. The social preferences R satisfy **O**. They also satisfy **RC** and **RWP**. To verify **HEF**, let ${}_1\mathbf{u}$, ${}_1\mathbf{v} \in X$ satisfy that $u_1 > v_1 > x > w$ and $u_t = w$ and $v_t = x$ for all $t > 1$. Then $W({}_1\mathbf{u}) = w$ and $W({}_1\mathbf{v}) = x$. Thus, $W({}_1\mathbf{v}) > W({}_1\mathbf{u})$, and so ${}_1\mathbf{v}\ P\ {}_1\mathbf{u}$. To verify **IF**, note that, for all ${}_1\mathbf{u} \in X$, $W({}_1\mathbf{u}) = W({}_2\mathbf{u})$. Hence, ${}_1\mathbf{u}\ R\ {}_1\mathbf{v}$ if and only if ${}_2\mathbf{u}\ R\ {}_2\mathbf{v}$ even if $u_1 = v_1$ does not hold. To see that R violates **WP**, note that ${}_1\mathbf{u}\ I\ {}_{\text{con}}0$ if ${}_1\mathbf{u} \in X$ is defined by, for all $t \geq 1$, $u_t = 1/t$. Clearly, R violates **STP**.

4.8 Concluding remarks

Condition **HEF** assigns priority to an infinite number of worse-off generations in comparisons where the assignment of such priority only reduces the utility of the better-off present generation. We consider this to be a compelling consequentialist equity condition. In particular, as discussed in section 4.3, the condition can be endorsed from both an egalitarian and a utilitarian point of view. It is therefore discouraging that condition **HEF** to such a large extent limits the possibility of being Paretian (cf. Propositions 2, 3, 4, and 5). In principle, there are two ways out of the ethical dilemma that these results pose.

One possibility is to drop continuity. In line with earlier literature, the analysis indicates that continuity conditions are not innocent technical assumptions; rather, such conditions have significant normative implications in the social evaluation of infinite utility streams (e.g., in the words of

Svensson, 1980, p. 1254, 'the continuity requirement is a value judgment'). By employing social preferences over infinite utility streams defined by Basu and Mitra (2007a), Asheim and Tungodden (2004b), and Bossert, Sprumont and Suzumura (2006) (and, if necessary, invoking Szpilrajn's (1930) Lemma to complete the preferences), we can establish the existence of two kinds of social preferences that satisfy **O**, **SP**, **HEF**, and **IF**: One is classical utilitarian, the other is egalitarian and based on leximin. Such preferences are appealing, since they satisfy 'Weak Anonymity' as well as the four consequentialist equity conditions listed in section 4.2. On the other hand, they are all insensitive toward the information provided by either interpersonal level comparability or interpersonal unit comparability. Classical utilitarianism makes no use of interpersonal level comparability (even if utilities are level comparable), while leximin makes no use of interpersonal unit comparability (even if utilities are unit comparable).

Another possibility is to weaken the Paretian requirement to condition **RWP**. Then, as reported in Example 3, there are social preferences satisfying **O**, **RC**, **HEF**, and **IF**. However, the social preferences presented in Example 3 are unappealing, since they entail invariance for the utility during any finite part of the stream. In particular, such social preferences do not satisfy Chichilnisky's (1996) 'No Dictatorship of the Future' condition. However, there are more attractive alternatives. It can be shown that conditions **O**, **RC**, **RWP**, **HEF**, and **IF** imply insensitivity for the interests of the present *only* when the present utility exceeds the stationary equivalent of the utility stream. The conditions do not preclude a trade-off between the interests of the present and future otherwise. Therefore, there exist social preferences satisfying conditions **O**, **RC**, **RWP**, **HEF**, and **IF** that are consistent with both of Chichilnisky's (1996) no-dictatorship conditions ('No Dictatorship of the Present' and 'No Dictatorship of the Future'), and make use of both interpersonal level comparability and interpersonal unit comparability of (at least) cardinally measurable fully comparable utilities. These possibilities are discussed in greater detail in Asheim and Tungodden (2006).

Thus, it is our view that the impossibility results reported in the present chapter should not be used to rule out 'Hammond Equity for the Future' and other consequentialist equity conditions as ethical guidelines for intergenerational equity. They do, however, show that consequentialist equity conditions seriously restrict the set of possible intergenerational social preferences.

References

Asheim, G.B. (1991) 'Unjust Intergenerational Allocations', *Journal of Economic Theory*, vol. 54(2), pp. 350–71.

Asheim, G.B., T. Mitra and B. Tungodden (2006) 'Sustainable Recursive Social Welfare Functions', Memorandum 18/2006, Department of Economics, University of Oslo.

Asheim, G.B. and B. Tungodden (2004a) 'Do Koopmans' Postulates Lead to Discounted Utilitarianism?', Discussion paper 32/04, Norwegian School of Economics and Business Administration.

Asheim, G.B. and B. Tungodden (2004b) 'Resolving Distributional Conflicts between Generations', *Economic Theory*, vol. 24(1), pp. 221–30.

Basu, K. and T. Mitra (2003) 'Aggregating Infinite Utility Streams with Intergenerational Equity: the Impossibility of Being Paretian', *Econometrica*, vol. 71(5), pp. 1557–63.

Basu, K. and T. Mitra (2007a) 'Utilitarianism for Infinite Utility Streams. A New Welfare Criterion and its Axiomatic Characterization', *Journal of Economic Theory*, vol. 133(1), pp. 350–73.

Basu, K. and T. Mitra (2007b) 'Possibility Theorems for Aggregating Infinite Utility Streams Equitably', Chapter 5 of this volume.

Birchenhall, C.R. and P. Grout (1979) 'On Equal Plans with an Infinite Horizon', *Journal of Economic Theory,* vol. 21(2), pp. 249–64.

Blackorby, C., D. Donaldson and J.A. Weymark (1984) 'Social Choice with Interpersonal Utility Comparisons: A Diagrammatic Introduction', *International Economic Review*, vol. 25(2), pp. 327–56.

Bossert, W., Y. Sprumont, and K. Suzumura (2006) 'Ordering Infinite Utility Streams', forthcoming in *Journal of Economic Theory*.

Chichilnisky, G. (1996) 'An Axiomatic Approach to Sustainable Development', *Social Choice and Welfare*, vol. 13(2), pp. 231–57.

Diamond, P. (1965) 'The Evaluation of Infinite Utility Streams', *Econometrica*, vol. 33(1), pp. 170–7.

Fields, G.S. and J.C.H. Fei (1978) 'On Inequality Comparisons', *Econometrica*, vol. 46(2), pp. 303–16.

Fleurbaey, M. and P. Michel (2001) 'Transfer Principles and Inequality Aversion, with Application to Optimal Growth', *Mathematical Social Sciences*, vol. 42(1), pp. 1–11.

Fleurbaey, M. and P. Michel (2003) 'Intertemporal Equity and the Extension of the Ramsey Principle', *Journal of Mathematical Economics*, vol. 39(7), pp. 777–802.

Hammond, P.J. (1976) 'Equity, Arrow's Conditions and Rawls' Difference Principle', *Econometrica*, vol. 44(4), pp. 793–804.

Koopmans, T.C. (1960) 'Stationary Ordinal Utility and Impatience', *Econometrica*, vol. 28(2), pp. 287–309.

Sakai, T. (2003) 'Intergenerational Preferences and Sensitivity to the Present', *Economics Bulletin*, vol. 4(26), pp. 1–6.

Sakai, T. (2006) 'Equitable Intergenerational Preferences on Restricted Domains', *Social Choice and Welfare*, vol. 27(1), pp. 41–54.

Shinotsuka, T. (1998) 'Equity, Continuity and Myopia: A Generalization of Diamond's Impossibility Theorem', *Social Choice and Welfare*, vol. 15(1), pp. 21–30.

Suzumura, K. and T. Shinotsuka (2003) 'On the possibility of continuous, Paretian and egalitarian evaluation of infinite utility streams', mimeo. Institute of Economic Research, Hitotsubashi University.

Svensson, L.-G. (1980) 'Equity among Generations', *Econometrica*, vol. 48(5), pp. 1251–6.

Szpilrajn, E. (1930) 'Sur l'extension du l'ordre partiel', *Fundamenta Mathematicae*, vol. 16, pp. 386–9.

Xu, Y. (2007) 'Pareto principle and intergenerational equity: Immediate impatience, universal indifference, and impossibility', Ch. 7 of this volume.

5
Possibility Theorems for Equitably Aggregating Infinite Utility Streams*

Kaushik Basu and Tapan Mitra

5.1 Introduction

The need to aggregate and evaluate infinite streams of returns or utility arises in several areas of economics, ranging from intergenerational welfare theory to environmental economics. The subject of intergenerational equity in the context of aggregating infinite utility streams has been of enduring interest to economists, starting with the work of Ramsey (1928), who had maintained that discounting one generation's utility or income vis-à-vis another's to be 'ethically indefensible', and something that 'arises merely from the weakness of the imagination.' His conjecture about the difficulty of aggregating infinite streams, while respecting intergenerational equity, turned out to be compelling, as a large number of impossibility theorems were proved subsequently by a number of authors, starting with the seminal works of Koopmans (1960) and Diamond (1965).

This problem has been confronted in the philosophy literature as well. Cowen and Parfit (1992), for instance, discussed the problem of aggregating the welfares of future generations, at length, and reached the conclusion that discounting the costs and benefits of future generations cannot be ethically justified. Hence, if we want to be morally correct, we must be 'against the discount rate'. The problem that they do not address and is germane to our chapter, is the logical feasibility of what they recommend. If we do decide to go along with their advice and give equal importance to future returns, then how do we aggregate future streams of returns when these stretch into infinity? Simply adding up will often not work – it may not give us a real number and could lead to a violation of the Pareto principle.

* This paper benefited greatly from a presentation and discussion at the IEA Roundtable Meeting in Hakone, Japan, 10–12 March 2005, and from a detailed comment by Yongsheng Xu. This is an area where, over the years, we have had conversations with and comments from a large number of economists and would like to thank, in particular, Geir Asheim, Claude d'Aspremont, Kuntal Banerjee, Marc Fleurbaey, Nick Kiefer, Wlodek Rabinowicz, Tomoichi Shinotsuka, Kotaro Suzumura and Jorgen Weibull.

Yet it would be wrong to abandon the effort to search for a social welfare function that aggregates infinite streams of returns and satisfies intergenerational anonymity and some form of the Pareto criterion. In reality, we encounter this problem all the time. In deciding whether to build a dam on a river, which will help irrigation and generate electricity but damage fauna and flora, we clearly face a problem of choosing between long streams of utility, stretching far into the future. Even if we believe that the world has a finite future, since we do not know its termination date, we effectively face an infinite decision problem.

Moreover, every time we analyze an infinitely repeated game, we are forced to confront an infinite decision problem. And, if we are to pass judgment on which among a set of possible outcomes is superior, we are compelled to contend with precisely the problem that is the concern of this chapter.

In Diamond's celebrated paper (1965) he had shown that there is no social welfare function that aggregates infinite utility streams while satisfying the Pareto condition, a weak form of anonymity and a continuity property.[1] In a recent paper (Basu and Mitra, 2003), we tried to show that the problem is more discouraging because the impossibility result survives even if we do not impose any continuity restriction on the social welfare function. Are we then completely into a *cul-de-sac*? This chapter tries to answer this in the negative.

We can think of many routes to getting possibility results. In an elegant paper, Svensson (1980) had shown that if, instead of seeking a (real-valued) social welfare function, we merely searched for the ability to rank infinite streams of utilities, then it is possible to prove that the requirements of equity and the Pareto principle are compatible. He does this, however, with the use of Szpilrajn's theorem, which implies a non-constructive proof. Related results have been obtained by Suzumura and Shinotsuka (2003) and Bossert, Sprumont and Suzumura (2004).

Though we delve briefly into this, our main aim in this study is to look for possibility theorems that satisfy *representability*; that is, the existence of real-valued social welfare functions. More precisely, our aim is to delineate the frontier of possibility and impossibility results for the existence of real-value social welfare functions. We consider, in particular, weakening the Pareto axiom and exploring domain restrictions.

It does seem that in reality the domain of values that individual utilities can take is often quite limited. The simple assumption that an individual's utility can be represented by any real number may be mathematically convenient, but it is unrealistic. Given the limits of human perception, it is much more realistic to suppose that individual utilities can take a finite number of values or, at most, a countably infinite number of values. Thus, exploring the implications of such domain restrictions certainly seems worthwhile.

Of course, domain restrictions by themselves will not yield possibility theorems, given the general impossibility theorem of Basu and Mitra (2003, Theorem 1), which applies to all domains, however restrictive they may be.[2] But, we try to show that, as soon as we combine domain restrictions with weaker versions of the Pareto axiom, the scope for the use of social welfare functions expands considerably (Theorem 3).

Our investigation also reveals that the particular nature of the domain restriction may be quite important for such possibility results. Under domain restrictions of other types, even the Weak Pareto axiom is seen to be incompatible with the requirement of an equitable social welfare function (Theorem 4). However, if the postulated version of Pareto is sufficiently weak, then it is possible to generate equitable and Paretian social welfare functions without any domain restrictions (Theorem 5).

It is true that the exercise that we undertake in this chapter is abstract and theoretical but it is motivated by the practical concern for shedding light on what is feasible once we reject the standard (inequitable) method of aggregating streams by discounting the returns that accrue to future generations.

5.2 Formal setting and basic results

Let $\mathbb{R}$ be the set of real numbers, $\mathbb{N}$ the set of positive integers, and $\mathbb{M}$ the set of non-negative integers. Suppose $Y \subset \mathbb{R}$ is the set of all possible utilities that any generation can achieve. Then $X = Y^{\mathbb{N}}$ is the set of all possible utility streams. If $\{x_t\} \in X$, then $\{x_t\} = (x_1, x_2, \ldots)$, where, for all $t \in \mathbb{N}$, $x_t \in Y$ represents the amount of utility that the generation of period t earns. For all $y, z \in X$, we write $y \geq z$ if $y_i \geq z_i$, for all $i \in \mathbb{N}$; we write $y > z$ if $y \geq z$ and $y \neq z$; and we write $y \gg z$, if $y_i > z_i$, for all $i \in N$.

If Y has only one element, then X is a singleton, and the problem of ranking or evaluating infinite utility streams is trivial. Thus, without further mention, the set Y will always be assumed to have at least two distinct elements.

A social welfare function (SWF) is a mapping $W : X \to \mathbb{R}$. Consider now the axioms that we may want the SWF to satisfy. The first axiom is the standard Pareto condition.

Pareto Axiom: For all $x, y \in X$, if $x > y$, then $W(x) > W(y)$.

The next axiom is the one that captures the notion of 'inter-generational equity'. We shall call it the 'anonymity axiom'.[3] It is equivalent to the notion of 'finite equitableness' (Svensson, 1980) or 'finite anonymity' (Basu, 1994).[4]

Anonymity Axiom: For all $x, y \in X$, if there exist $i, j \in \mathbb{N}$ such that $x_i = y_j$ and $x_j = y_i$, and for every $k \in \mathbb{N} \sim \{i, j\}$, $x_k = y_k$, then $W(x) = W(y)$.

We shall begin by stating the main impossibility theorem that was established in Basu and Mitra (2003, Theorem 1). This will be the setting in which we can then ask the question of what is possible.

Theorem 1 *There does not exist any SWF satisfying the Pareto and Anonymity Axioms.*

It is the rather sparse requirement of this theorem that is at the root of the frustration that this field of inquiry has generated. Note, in particular, that the impossibility result does not depend on any continuity postulate on the SWF; and, it applies to all domains of the SWF.

Before exploring the routes out of this, it is useful to place the problem in perspective by recalling Svensson's (1980) important theorem. Let us suppose that we abandon the search for an SWF and instead look for a social welfare ordering[5] (SWO). We then have the result due to Svensson (1980) that there is an SWO which satisfies the (appropriate relational versions of the) Pareto and Anonymity axioms. For reasons of completeness we briefly review Svensson's result. We do this also because the use of a variant of Szpilrajn's Theorem (due to Suzumura, 1983, Theorem A(5)) allows us to give a particularly easy proof of it. Furthermore, Svensson (1980) restricts his exercise to the case where Y is the closed interval [0,1]; we state the version of his result which applies to any utility space Y. His proof, as well as ours, applies to this more general setting.

Formally, an SWO is a binary relation, $\succcurlyeq$ on X, which is complete and transitive. We use $\succ$ and $\sim$ to denote, respectively, the asymmetric and symmetric parts of $\succcurlyeq$. The properties of Pareto and Anonymity for an SWO are easy to define. We shall call these axioms $\succcurlyeq$-Pareto and $\succcurlyeq$-Anonymity to distinguish them from the axioms applied to an SWF.

$\succcurlyeq$-Pareto Axiom: For all $x, y \in X$, $x > y$ implies $x \succ y$.

$\succcurlyeq$-Anonymity Axiom: For all $x, y \in X$, if there exist $i, j \in \mathbb{N}$, such that $x_i = y_j$ and $x_j = y_i$ and for every $k \in \mathbb{N} \sim \{i, j\}$, $x_k = y_k$, then $x \sim y$.

First, let us give a statement of Suzumura's result. Let Ω be a set of alternatives. If R is a binary relation on Ω and R^* an ordering on Ω, we shall say that R^* is an *ordering extension* of R if, for all $x, y \in \Omega$, xRy implies xR^*y. We say that R is *consistent* if, for all $t \in \mathbb{N}$, and for all $x^1, x^2, \ldots, x^t \in \Omega$, $[x^1Rx^2$ and not x^2Rx^1, and for all $k \in \{2, 3, \ldots, t-1\}$, $x^kRx^{k+1}]$ implies not x^tRx^1.

Lemma 1 (Szpilrajn's Corollary [Suzumura, 1983]): *A binary relation R on Ω has an ordering extension if and only if it is consistent.*

Before proving the next theorem it is useful to introduce some new notation. If $\sigma : \mathbb{N} \to \mathbb{N}$ is a permutation, and there exists $t \in \mathbb{N}$, such that for all $k > t$, $\sigma(k) = k$, then we shall call σ a *finite permutation*. Given a finite permutation, σ, we shall use $n(\sigma)$ to denote the smallest integer t which has the property that, for all $k > t$, $\sigma(k) = k$. Given a finite permutation, σ, and $x \in X$, we shall use $x(\sigma)$ to denote $y \in X$, where y is obtained by permuting the elements of x using σ.

In contrast to Theorem 1, we now have:

Theorem 2 (Svensson, 1980): *There exists a social welfare ordering satisfying the $\succcurlyeq$-Pareto and $\succcurlyeq$-Anonymity Axioms.*

Proof. Define two binary relations, P and I, on X, as follows. For all $x, y \in X$, if $x > y$ then xPy. And if there exists i, j such that $x_i = y_j$ and $x_j = y_i$, and $x_k = y_k$ for all $k \neq i, j$, then xIy. Now define the binary relation R as follows: $xRy \Leftrightarrow xPy$ or xIy.

To see that R is consistent, suppose $t \in N$ and $x^1, x^2, \ldots, x^t \in X$ such that

(A) x^1Rx^2 and not x^2Rx^1, and
(B) x^kRx^{k+1} for all $k \in \{2, 3, \ldots t-1\}$.

We have to show that not x^tRx^1.

Note that (A) and (B) can be written equivalently as

(A′) x^1Px^2, and
(B′) x^kPx^{k+1} or x^kIx^{k+1} for all $k \in \{2, 3, \ldots t-1\}$.

Note that (A′) and [x^2Px^3 or x^2Ix^3] imply that there exists a finite permutation, σ_3, such that:

(A″) $x^1Px^3(\sigma_3)$.

Next note that (A″) and [x^3Px^4 or x^3Ix^4] imply that there exists a finite permutation, σ_4, such that: $x^1Px^4(\sigma_4)$. Continuing in the same way we get the result that there exists a finite permutation, σ_t, such that: $x^1Px^t(\sigma_t)$. This implies not [x^tPx^1 or x^tIx^1]. Therefore, not x^tRx^1.

Hence, by Szpilrajn's Corollary, R has an ordering extension $\succcurlyeq$. Clearly $\succcurlyeq$ satisfies the $\succcurlyeq$-Pareto Axiom and the $\succcurlyeq$-Anonymity Axiom.

For a long time, researchers have conjectured that the impossibility of having a social welfare function satisfying Pareto and anonymity was a problem of *representability*; that is, of there not being 'enough real numbers' to do the job. Since Diamond's theorem (1965) showed that the requirements of Pareto, anonymity *and continuity* were inconsistent, the conjecture remained an open one. But in the light of Theorem 1 above we can state a corollary which (a) confirms the conjecture, and (b) clarifies the relation between Theorems 1 and 2 in a way that is especially useful. Toward this end, define:

Representability: A SWO, $\succcurlyeq$, is *representable* if there exists a mapping, $f : X \to \mathbb{R}$ such that, for all $x, y \in X$, $x \succcurlyeq y \Leftrightarrow f(x) \geq f(y)$.

In the light of Svensson's result, Theorem 1 can be restated as follows.

Corollary 1 *There does not exist a SWO satisfying the $\succcurlyeq$-Pareto Axiom, the $\succcurlyeq$-Anonymity Axiom and representability.*

Proof. If a representable SWO satisfies the $\succcurlyeq$-Pareto Axiom and the $\succcurlyeq$-Anonymity Axiom, the real-valued function, $f : X \to \mathbb{R}$ that represents the SWO, must satisfy the Pareto and Anonymity Axioms. But we know from Theorem 1 that no such f exists. This establishes the result.

Corollary 1 makes the nature of the impossibility clear. If we are looking for an equitable SWO (that is, one satisfying the anonymity principle) to evaluate infinite streams of returns, we have to be prepared to weaken the Pareto axiom or to give up the representability requirement. There is a case for exploring both these avenues. In a recent paper Bossert, Sprumont and Suzumura (2004) have looked at the possibilities that emerge when one does not require representability.[6] In what follows, we explore what is possible by relaxing the Pareto axiom.

5.3 Weakening Pareto

It is arguable that for certain philosophical and even policy purposes we do not need the full power of the Pareto condition (even if we are committed Paretians) simply because all the possibilities that are technically allowed in our specification of the domain may not arise under any eventuality. Indeed for certain ethical discourses involving the comparison of the moral worth of individual actions and universalizable rules (see Basu (1994)) it may be enough to be armed with some weaker forms of Paretianism.

One idea that may be of interest is to restrict the analysis to cases where one state is obtained from another through changes in a finite number of periods. For such cases it is enough to use the following weakening of Pareto that we shall call 'weak dominance.'

Weak Dominance Axiom: For all $x, y \in X$, if for some $j \in \mathbb{N}$, $x_j > y_j$, while, for all $k \neq j$, $x_k = y_k$, then $W(x) > W(y)$.

Another version of Pareto – this one has been widely used in the literature (see Arrow, 1963; Sen, 1977) – is the 'Weak Pareto' axiom, as defined below.[7]

Weak Pareto Axiom: For all $x, y \in X$, if $x \gg y$, then $W(x) > W(y)$.

A natural next step is to consider an axiom that combines the two above axioms. That is precisely what the next axiom does.

Partial Pareto Axiom: The SWF, W, satisfies the Weak Dominance axiom and the Weak Pareto Axiom.

The Partial Pareto Axiom demands that the SWF be positively sensitive to an increase in utility of a single generation, the utilities of other generations being unchanged (and therefore that it be positively sensitive to increases in utilities of any finite number of generations, the utilities of other generations being unchanged), and also that the SWF be positively sensitive to an increase in utilities of all generations. However, it need not be positively sensitive to an increase in utilities of an infinite number of generations, when the utilities of a (non-empty) set of generations is unchanged. This is the principal difference between the Partial Pareto axiom and the Pareto axiom.

Possibility results for restricted domains

Note that if we recognize that human perception or cognition is not endlessly fine, so that sufficiently small changes in well-being go unperceived, it seems reasonable to suppose that the set of feasible utilities will be a discrete set.[8] The same is true if the benefits are measured in money and there is a well-defined smallest unit, as is true for all currencies (Segerberg 1976). Thus, it seems worthwhile to explore whether, with $Y \subset \mathbb{M}$ (which captures this very reasonable possibility), there is a social welfare function (on X) respecting Anonymity and one of the weaker versions of the Pareto axiom, introduced above.[9] It is interesting to note that the domain restriction allows us to establish the existence of an equitable SWF, which satisfies the strongest of these versions of Pareto, namely the Partial Pareto axiom.

Proposition 1 *Assume $Y \subset \mathbb{M}$. There exists an SWF satisfying the Partial Pareto and Anonymity Axioms.*

Proof. For each $x \in X$, let $E(x) = \{y \in X$: there is some $N \in \mathbb{N}$, such that $y_k = x_k$ for all $k \in \mathbb{N}$, which are $\geq N\}$. Let $\Im$ be the collection $\{E : E = E(x)$ for some $x \in X\}$. Then $\Im$ is a partition of X. That is, if E and F belong to $\Im$, then either $E = F$, or E is disjoint from F; further, $\cup_{E \in \Im} E = X$.

Define a function, $f : X \to \mathbb{M}$ as follows. Given any $x \in X$, let $f(x) = \min\{x_1, x_2, \ldots\}$. Since $x_i \in \mathbb{M}$ for all $i \in \mathbb{N}$, the set $\{x_1, x_2, \ldots\}$ is a non-empty subset of the set of non-negative integers and therefore has a smallest element (Munkres, 1975, p. 32). Thus, f is well-defined. By the axiom of choice, there is a function, $g : \Im \to X$, such that $g(E) \in E$ for each $E \in \Im$.

Given any $x \in X$, we can denote for each $N \geq 1$, $(x_1, \ldots, x_N)$ by $x(N)$, and $(x_1 + \cdots + x_N)$ by $I(x(N))$. Next, given any x, y in $E \in \Im$, define $h(x, y) = \lim_{N \to \infty} [I(x(N)) - I(y(N))]$. Notice that h is well-defined, since given any x, y in $E \in \Im$, there is some $M \in \mathbb{N}$, such that $[I(x(N)) - I(y(N))]$ is a constant for all $N \geq M$. Given any x, y in $E \in \Im$, define $H(x, y) = 0.5[h(x, y)/[1 + |h(x, y)|]]$. Then $H(x, y) \in (-0.5, 0.5)$.

We now define $W : X \to \mathbb{R}$ as follows. Given any $x \in X$, we associate with it its equivalence class, $E(x)$. Then, using the function g, we get $g(E(x)) \in E(x)$. Next, using the functions, h and H, we obtain $h(x, g(E(x)))$ and $H(x, g(E(x)))$. Finally, define $W(x) = f(x) + H(x, g(E(x)))$.

The Anonymity Axiom can be verified as follows. If x, y are in X, and there exist i, j in $\mathbb{N}$, such that $x_i = y_j$ and $x_j = y_i$, while $x_k = y_k$ for all $k \in \mathbb{N}$, such that $k \neq i, j$, then $E(x) = E(y)$. Furthermore, denoting this common set by E, we see that $h(x, g(E)) = h(y, g(E))$, and so $H(x, g(E)) = H(y, g(E))$. Further, the set $\{x_1, x_2, \ldots\}$ is the same as the set $\{y_1, y_2, \ldots\}$, so that $f(x) = f(y)$. Thus, we obtain: $W(x) = W(y)$.

The Partial Pareto Axiom can be verified as follows. If x, y are in X, and there exists $i \in \mathbb{N}$, such that $x_i > y_i$, while $x_k = y_k$ for all $k \in \mathbb{N}$, such that $k \neq i$, then $E(x) = E(y)$. Furthermore, denoting the common set by E, we see that $h(x, g(E)) > h(y, g(E))$. This implies $H(x, g(E)) > H(y, g(E))$. Further, the smallest element of the set $\{x_1, x_2, \ldots\}$ is at least as large as the smallest element of the set $\{y_1, y_2, \ldots\}$, so that we have $f(x) \geq f(y)$. Thus, we obtain the desired inequality: $W(x) > W(y)$.

If $x, y \in X$, and $x \gg y$, then $E(x) \neq E(y)$. Thus, we will not be able to compare $H(x, g(E(X)))$ with $H(y, g(E(y)))$. However, we do know that $H(x, g(E(x))) > -0.5$, and $H(y, g(E(y))) < 0.5$. Further, since $x \gg y$, we have $f(x) \geq f(y) + 1$. Thus, we obtain:

$$W(x) = f(x) + H(x, g(E)) > f(y) + 1 - 0.5 > f(y) + H(y, g(E)) = W(y).$$

Proposition 1 has two shortcomings. First, it is a possibility result for a social welfare function, but we do not know how to construct the social welfare function whose existence is asserted, since our proof uses the Axiom of Choice.[10] The possible policy use of Proposition 1 is therefore limited. We should clarify, however, that though we give a proof using the axiom of choice and indeed know of no other proof, it is not the case that we have proved that the axiom of choice is necessary. Indeed, it remains a bit of an open conjecture as to whether the axiom of choice is *necessary* for the above proposition.

The second shortcoming can be seen by considering the set-up, where $Y = \{0, 1\}$, so that we have the strongest possible domain restriction. Theorem 1 implies that there is no SWF respecting the Pareto and Anonymity Axioms. And, Proposition 1 implies that there is an SWF satisfying the Partial Pareto and Anonymity Axioms. It follows that any social welfare function, W, so obtained, must violate the Pareto principle in a way that is particularly disturbing; that is, it must be the case that there exist alternatives $x, y \in X$ such that $x > y$, but $W(x) < W(y)$.

To see this, suppose on the contrary that there is an SWF, W, satisfying the Anonymity and Partial Pareto axioms, and the following monotonicity condition:

Condition M (*Monotonicity*) For all $x, y \in X$, if $x > y$, then $W(x) \geq W(y)$.

We claim then that W must, in fact, satisfy the Pareto Axiom. To see this, let $x, y \in X$ with $x > y$. There are three possibilities: (i) $x \gg y$ (ii) $x_i > y_i$ for $i \in F$, where F is a finite subset of $\mathbb{N}$ and $x_i = y_i$ for all $i \in \mathbb{N} \sim F$, (iii) $x_i > y_i$ for $i \in I$, where I is an infinite strict subset of $\mathbb{N}$, and $x_i = y_i$ for all $i \in \mathbb{N} \sim I$. In cases (i) and (ii), by the Partial Pareto axiom, we must have $W(x) > W(y)$. In case (iii), let j be the smallest index in I, and define z by $z_j = y_j$ and $z_i = x_i$ for all $i \neq j$.

Then, $z \in X$, and $z > y$, so that by Condition M, $W(z) \geq W(y)$. Also, comparing x and z, we see that they differ in only the j-th index, and $x_j > y_j = z_j$, so that the Partial Pareto axiom implies that $W(x) > W(z)$. Thus, $W(x) > W(y)$, and our claim is established. But, by Theorem 1, there is no SWF satisfying the Pareto and Anonymity axioms. Consequently, any SWF, W, satisfying the Anonymity and Partial Pareto axioms, must violate Condition M.[11]

Both the shortcomings of Proposition 1 arise from the fact that we are trying to define a social welfare function, which is sensitive to the utility of a single generation, when the utilities of all other generations are unchanged. If we give up this sensitivity, and weaken our Partial Pareto requirement to the Weak Pareto one, we get a particularly satisfying possibility result on all domains X, when $Y \subset \mathbb{M}$.

Theorem 3 *Assume $Y \subset \mathbb{M}$. Then the SWF, $W : X \to \mathbb{M}$, given by*:

$$W(x) = \min\{x_1, x_2, \ldots\} \text{ for all } x \in X$$

satisfies the Weak Pareto and Anonymity Axioms. Further, it satisfies Condition M.

Proof. The function, $W : X \to \mathbb{M}$, given by $W(x) = \min\{x_1, x_2, \ldots\}$ for all $x \in X$, is well-defined (as already noted in the proof of Proposition 1). If $x, y \in X$ and $x \gg y$, then denoting an index, for which $\min\{x_1, x_2, \ldots\}$ is attained, by $k \in \mathbb{N}$ we have:

$$W(y) = \min\{y_1, y_2, \ldots\} \leq y_k < x_k = \min\{x_1, x_2, \ldots\} = W(x)$$

so that the Weak Pareto axiom is satisfied.

If $x, y \in X$, and there exist $i, j \in \mathbb{N}$, such that $x_i = y_j$ and $x_j = y_i$, while $x_k = y_k$ for all $k \in \mathbb{N}$, such that $k \neq i, j$, then the set $\{x_1, x_2, \ldots\}$ is the same as the set $\{y_1, y_2, \ldots\}$, so that $W(x) = W(y)$. Thus, the Anonymity axiom is satisfied.

Finally, if $x, y \in X$ and $x > y$, then denoting an index, for which $\min\{x_1, x_2, \ldots\}$ is attained, by $k \in \mathbb{N}$, we have:

$$W(y) = \min\{y_1, y_2, \ldots\} \leq y_k \leq x_k = \min\{x_1, x_2, \ldots\} = W(x)$$

so that Condition M is satisfied.

The social welfare function in Theorem 3 can be explicitly written down, and this makes the possibility result especially useful for policy purposes.

Weakening domain restrictions

The above possibility results are obtained by weakening the Pareto axiom (to Partial Pareto or to Weak Pareto) and also considering a discrete domain. How would a change in the latter affect the results? It is especially useful to ask this question in the context where $Y = [0, 1]$, since this is the standard framework used by Koopmans (1960), Diamond (1965), Svensson (1980) and others.

As it turns out, we run again into impossibility results, which means that with $Y=[0,1]$, the weakening of Pareto to Partial Pareto or to Weak Pareto does not help to reverse the impossibility result of Theorem 1. To establish the first of these impossibility results, which follows directly from the result of Basu and Mitra (2003, Theorem 2), it is useful to introduce a new axiom, the interest in which is purely constructive, so as to be able to explain the next result clearly.

Dominance Axiom: For all $x, y \in X$, if there exists $j \in \mathbb{N}$ such that $x_j > y_j$, and, for all $k \neq j$, $x_k = y_k$, then $W(x) > W(y)$. For all $x, y \in X$, if $x \gg y$, then $W(x) \geq W(y)$.

Note that the last inequality in the statement of this axiom is a weak inequality, unlike in the definition of the Partial Pareto Axiom. Hence, Partial Pareto is stronger than Dominance (which in turn is stronger than Weak Dominance).[12]

Proposition 2 *Assume* $Y \supset [0,1]$. *There is no SWF satisfying the Partial Pareto and Anonymity Axioms.*

Proof. By Theorem 2 of Basu and Mitra (2003), we know that there is no SWF satisfying the Dominance and Anonymity axioms. The result is proved by noting that the Partial Pareto axiom is stronger than the Dominance axiom.

When we weaken the Partial Pareto axiom (of Proposition 2) to Weak Pareto, the impossibility result persists, but it is a more subtle result, since the sensitivity of the SWF to a change in a single generation's utility (when the utilities of all other generations are unchanged) is not being imposed. The proof of it is, likewise, more intricate, combining the methods used by Basu and Mitra (2003, Theorem 2) and by Fleurbaey and Michel (2003).

Theorem 4 *Assume* $Y \supset [0,1]$. *There is no SWF satisfying the Weak Pareto Axiom and the Anonymity Axiom.*

Proof. To establish the theorem, assume that there exists a social welfare function, $W: X \to \mathbb{R}$, which satisfies the Weak Pareto and Anonymity Axioms.

Denote the vector (1,1,1,...) in X by e. Define the sequences x and y in X as follows:

$$x = \left(\frac{1}{4}, \frac{2}{4}, \frac{3}{4}, \ldots, \frac{1}{4^k}, \ldots, \frac{4^k-1}{4^k}, \ldots\right) \tag{5.1}$$

$$y = \left(\frac{1}{4} + \frac{1}{16}, \frac{2}{4} + \frac{1}{16}, \frac{3}{4} + \frac{1}{16}, \ldots, \frac{1}{4^k} + \frac{1}{4^{k+1}}, \ldots, \frac{4^k-1}{4^k} + \frac{1}{4^{k+1}}, \ldots\right) \tag{5.2}$$

For $s \in I \equiv (-0.5, 0.5)$, define:

$$y(s) = 0.5y + 0.25(1+s)e \tag{5.3}$$

Then $(1/8)e \leq y(s) \leq (7/8)e$, and so $y(s) \in X$ for each $s \in I$.

Define the function, $f : I \to \mathbb{R}$ by: $f(s) = W(y(s))$. By the Weak Pareto Axiom, f is monotonic increasing in s on I. Thus f has only a countable number of points of discontinuity in I. Let $a \in I$ be a point of continuity of the function f.

Define the sequence $x(a)$ as follows:

$$x(a) = 0.5x + 0.25(1+a)e \tag{5.4}$$

Clearly, $x(a) \in X$ and $y(a) \gg x(a)$. By the Weak Pareto Axiom, $W(y(a)) > W(x(a))$. We denote $[W(y(a)) - W(x(a))]$ by θ; then $\theta > 0$.

Denote $\max(0.5 - a, 0.5 + a)$ by Δ; then, $\Delta > 0$. Since f is continuous at a, given the θ defined above, there exists $\delta \in (0, \Delta)$, such that: $0 < |s - a| < \delta$ implies $|f(s) - f(a)| < \theta$. Note that for $0 < |s - a| < \delta$, we always have $s \in I$.

For $p \in \mathbb{N}$, let $r(p)$ denote the first non-zero remainder of the successive divisions of p by 4, and $q(p)$ the number of divisions with a zero remainder. (For example, $r(52) = 1$ and $q(52) = 1$.)

Define (following Fleurbaey and Michel (2003, p. 796)), for each $k \in \mathbb{N}$, a sequence x^k as follows:

$$x^k = \left(\frac{1}{4} + \frac{1}{16}, \frac{2}{4} + \frac{1}{16}, \frac{3}{4} + \frac{1}{16}, \ldots, \frac{1}{4^k} + \frac{1}{4^{k+1}}, \ldots, \frac{4^k - 1}{4^k} + \frac{1}{4^{k+1}}, \right.$$
$$\frac{1}{4^{k+1}}, \ldots, \frac{4p}{4^{k+1}}, \ldots, \frac{4p}{4^{k+1}}, \frac{4p+2}{4^{k+1}}, \frac{4p+3}{4^{k+1}}, \ldots, \frac{4^{k+1} - 1}{4^{k+1}}, \frac{1}{4^{k+2}},$$
$$\left. \frac{2}{4^{k+2}}, \ldots, \frac{4^{k+2} - 1}{4^{k+2}}, \ldots \right) \tag{5.5}$$

where p runs from 1 to $4^k - 1$, and the term $[4p/(4^{k+1})]$ is repeated $q(4p)$ times if $r(4p) = 1$, and $q(4p) + 1$ times otherwise. Now, for each $k \in \mathbb{N}$, we use x^k to define $x^k(a)$ as follows:

$$x^k(a) = 0.5x^k + 0.25(1+a)e \tag{5.6}$$

Clearly, $x^k(a) \in X$ for each $k \in \mathbb{N}$. Comparing the expressions for $x(a)$ and $x^1(a)$ in (5.4) and (5.6) respectively, we see that $x^1(a)$ is obtained from $x(a)$ by a finite permutation, and that for all $k > 1$, $x^k(a)$ is obtained from $x^{k-1}(a)$ by a finite permutation. Thus, for every $k \in \mathbb{N}$, $x^k(a)$ is obtained from $x(a)$ by a finite permutation, and the Anonymity Axiom yields:

$$W(x^k(a)) = W(x(a)) \quad \text{for all } k \in \mathbb{N} \tag{5.7}$$

Choose $K \in \mathrm{N}$ with $K \geq 2$ such that $(1/4^{K-2}) < \delta$, and define $S = (a - (1/4^{K-2}))$. We note that $0 < (a - S) < \delta$, and so $S \in I$, and:

$$W(y(S)) = f(S) > f(a) - \theta = W(y(a)) - \theta \tag{5.8}$$

We now compare the welfare levels associated with $x^K(a)$ and $y(S)$ as follows. Notice that:

$$\begin{aligned}
x^K(a) &= 0.5x^K + 0.25(1+a)e = 0.5y + 0.25(1+a)e - 0.5(y - x^K)\\
&= y(a) - 0.5(y - x^K)\\
&\geq y(a) - 0.5(1/4^K)e\\
&= 0.5y + 0.25(1+a)e - 0.5(1/4^K)e\\
&> 0.5y + 0.25(1 + a - (1/4^{K-1}))e\\
&= 0.5y + 0.25(1 + a - (1/4^{K-2}))e + 0.25(3/4^{K-1})e\\
&\gg 0.5y + 0.25(1+S)e = y(S)
\end{aligned}$$

Thus, by the Weak Pareto Axiom, we have:

$$W(x^K(a)) > W(y(S)) \tag{5.9}$$

Using (5.7), (5.8) and (5.9), we obtain:

$$\begin{aligned}
W(y(a)) - \theta &= W(x(a))\\
&= W(x^K(a))\\
&> W(y(S))\\
&> W(y(a)) - \theta
\end{aligned}$$

a contradiction, which establishes our result.

It is worth noting that, with the domain restriction $Y \subset \mathbb{M}$, weakening the Pareto axiom to the Weak Pareto axiom led to a reversal of the impossibility result of Theorem 1 to the possibility result of Theorem 3. When $Y = [0, 1]$, a similar weakening of the Pareto axiom (to the Weak Pareto axiom) does *not* produce such a reversal.

This suggests that to recover possibility when $Y = [0, 1]$, we need to go to a weaker form of Pareto. In fact, Weak Dominance is *not* weaker than Weak Pareto, but we *can* establish the existence of an equitable SWF, which satisfies

Weak Dominance. In fact, this possibility result holds with *no* domain restriction. Our proof employs the idea, already used in the proof of Proposition 1, of partitioning X into sets such that the members of each set differ from each other in only a finite number of indices. The proof of the possibility result then crucially hinges on (i) the use of the Axiom of Choice, and (ii) the fact that Weak Dominance never requires one to compare the welfare of members in two different sets of the partition.

Theorem 5 *There exists an SWF satisfying the Weak Dominance and Anonymity Axioms.*

Proof. For each $x \in X$, let $E(x) = \{y \in X$: there is some $N \in \mathbb{N}$, such that $y_k = x_k$ for all $k \in \mathbb{N}$, which are $\geq N\}$. Let $\Im$ be the collection $\{E : E = E(x)$ for some $x \in X\}$. Then, $\Im$ is a partition of X. By the axiom of choice, there is a function, $g : \Im \to X$, such that $g(E) \in E$, for each $E \in \Im$.

Given any x, y in $E \in \Im$, define $h(x, y) = \lim_{N\to\infty} [I(x(N)) - I(y(N))]$. We now define $W : X \to \mathbb{R}$ as follows. Given any $x \in X$, we associate with it its equivalence class, $E(x)$. Then, using g, we get $g(E(x)) \in E(x)$, and, using h, we obtain $h(x, g(E(x)))$. Now, define $W(x) = h(x, g(E(x)))$. The Anonymity Axiom and the Weak Dominance Axioms are easily verified.

Remarks:
(i) The Weak Dominance Axiom compares utility streams which differ for only one generation. One could define a concept of Finite Dominance, which allows for comparisons between utility streams, in which one utility stream always has at least as much utility for each generation as the other and the utility streams differ for at most a *finite* number of generations.

Finite Dominance: If $x, y \in X$, and $x > y$, and there is $N \in \mathbb{N}$ such that $x_k = y_k$ for all $k > N$, then $W(x) > W(y)$.

Clearly, W satisfies Finite Dominance if and only if it satisfies Weak Dominance. In view of this, Theorem 5 is *equivalent* to the statement obtained by replacing 'Weak Dominance' by 'Finite Dominance'.

(ii) The possibility result of Theorem 5 can be contrasted with the impossibility result of Diamond (1965). When $Y = [0, 1]$, Diamond's result shows that there is no social welfare order, continuous in the sup metric, which satisfies the relational versions of the Anonymity and Pareto Axioms. However, the proof of Diamond's impossibility result can be used to infer that there is no social welfare order, continuous in the sup metric, which satisfies the relational versions of Anonymity and *Weak Dominance*. Thus, continuity in the sup metric (in conjunction with Anonymity) is a stronger restriction than *representability* of a social welfare order in this context.

(iii) It can be checked that any SWF, W, satisfying the Weak Dominance and Anonymity Conditions must violate Condition M; that is, there must exist $x, y \in X$, such that $x > y$ and $W(x) < W(y)$.

5.4 Concluding remarks

We wanted to demarcate the boundary between what is possible and what is not and the set of results established in this chapter tries to do that vis-à-vis variations of the Pareto Axiom and the domain restriction for utilities. In setting out to write this chapter we had wanted to display the positive side of this field, namely, the possibility theorems. We have done so. But now, at the chapter's end, we find that in the process we have also highlighted the robustness of the impossibility theorems of the literature. This is probably a reminder that we have no option but to play the hand that we are dealt.

Our investigation of domain restrictions for possibility theorems of equitable social welfare functions is, of course, not complete. If we restrict our attention to the Weak Pareto axiom, we have the possibility theorem (Theorem 3) when $Y = \mathbb{M}$ and the impossibility theorem (Theorem 4) when $Y = [0, 1]$. One might be interested to know what results would hold if the domain restrictions place Y somewhere 'in between' these two cases. For example, Y could be the set of rationals in $[0, 1]$. Neither the method used to establish the possibility theorem when $Y = \mathbb{M}$, nor the method used to prove the impossibility result when $Y = [0, 1]$, applies in this case.[13] We might hope that future research in this area will develop new methods capable of dealing with a wider class of domain restrictions.

Notes

1. The continuity property postulated by Diamond is with respect to the sup metric on $X = [0, 1]^{\mathbb{N}}$.
2. Of course, the case in which the period utility space is a singleton, and so the domain of the social welfare function is also a singleton, is ruled out in the framework of Basu and Mitra (2003, Theorem 1).
3. In informal discussions throughout the chapter, the terms 'equity' and 'anonymity' are used interchangeably.
4. The Anonymity Axiom figures prominently in the social choice theory literature, where it is stated as follows: the social ordering is invariant to the information regarding individual orderings as to who holds which preference ordering. Thus, interchanging individual preference profiles does not change the social preference profile. For discussions of this axiom and its acceptability see May (1952) and Sen (1970, 1977).
5. An ordering is a binary relation which is complete and transitive.
6. In this connection, see also the papers by Suzumura and Shinotsuka (2003) and Xu (2005).
7. In fact, in some of the literature, what we are calling 'Weak Pareto' is often called 'Pareto', with the suffix 'strong' added to what we have called simply the 'Pareto axiom'.

8. The idea of setting a limit to the fineness of human perception has been used in a different context by Armstrong (1939) to argue that it is unreasonable to suppose that indifference is a transitive relation. For a discussion of this issue in individual choice theory, see Majumdar (1962).
9. While our choice of Y as a subset of the set of non-negative integers is motivated by the imprecision of human perception, the mathematical technique used to obtain our possibility result applies also to the case where $Y = \{(1/n) : n \in \mathbb{N}\}$, where clearly human perception has to be considered to be sufficiently refined.
10. The use of the Axiom of Choice in proving *impossibility results* is, perhaps, less objectionable.
11. A weak version of Pareto, which requires that Condition M, together with what we have called the Weak Pareto axiom, be satisfied, is quite appealing, and has been proposed and examined by Diamond (1965).
12. It is also worth noting that between Dominance and Weak Pareto, neither is stronger than the other. They are in fact non-comparable in terms of strength. The same is true between Weak Dominance and Weak Pareto.
13. Of course, since the streams considered in the proof of Theorem 4 consist entirely of rational entries, imposing a continuity (in the sup metric) axiom on the social welfare function will provide an impossibility result. But, using such a continuity axiom goes against the spirit of our chapter.

References

Armstrong, W.E. (1939) 'The Determinateness of the Utility Function', *Economic Journal*, vol. 49, pp. 453–67.

Arrow, K.J. (1963) *Social Choice and Individual Values*, 2nd edn (New York: John Wiley).

Basu, K. (1994) 'Group Rationality, Utilitarianism and Escher's "Waterfall"', *Games and Economic Behavior*, vol. 7, pp. 1–9.

Basu, K. and T. Mitra (2003) 'Aggregating Infinite Utility Streams with Intergenerational Equity: The Impossibility of Being Paretian', *Econometrica*, vol. 79, pp. 1557–63.

Bossert, W., Y. Sprumont and K. Suzumura (2004) 'The Possibility of Ordering Infinite Utility Streams', mimeo: Université de Montreal.

Cowen, T. and D. Parfit (1992) 'Against the Social Discount Rate', in P. Laslett and J.S. Fishkin (eds), *Justice between Age-Groups and Generations* (New Haven: Yale University Press).

Diamond, P. (1965) 'The Evaluation of Infinite Utility Streams', *Econometrica*, vol. 33, pp. 170–7.

Fleurbaey, M. and P. Michel (2003) 'Intertemporal Equity and the Extension of the Ramsey Principle', *Journal of Mathematical Economics*, vol. 39, pp. 777–802.

Koopmans, T.C. (1960) 'Stationary Ordinal Utility and Impatience', *Econometrica*, vol. 28, pp. 287–309.

Majumdar, T. (1962) *The Measurement of Utility*, 2nd edn (London: Macmillan).

May, K.O. (1952) 'A Set of Independent Necessary and Sufficient Conditions for Simple Majority Decision', *Econometrica*, vol. 20, pp. 680–4.

Munkres, J. (1975) *Topology* (London: Prentice Hall).

Ramsey, F.P. (1928) 'A Mathematical Theory of Savings', *Economic Journal*, vol. 38, pp. 543–59.

Segerberg, K. (1976) 'A Neglected Family of Aggregation Problems in Ethics', *Nous*, vol. 10, pp. 221–44.
Sen, A.K. (1970) *Collective Choice and Social Welfare* (San Francisco: Holden-Day).
Sen, A.K. (1977) 'On Weights and Measures: Informational Constraints in Social Welfare Analysis', *Econometrica*, vol. 45, pp. 1539–72.
Suzumura, K. (1983) *Rational Choice, Collective Decisions and Social Welfare* (Cambridge: Cambridge University Press).
Suzumura, K. and T. Shinotsuka (2003) 'On the Possibility of Continuous, Paretian and Egalitarian Evaluation of Infinite Utility Streams', mimeo: Institute of Economic Research, Hitotsubashi University.
Svensson, L.-G. (1980) 'Equity among Generations', *Econometrica*, vol. 48, pp. 1251–6.
Xu, Y. (2007) 'Pareto Principle and Intergenerational Equity: Immediate Impatience, Universal Indifference and Impossibility', Ch. 7 of this volume.

6

On the Existence of Paretian Social Welfare Quasi-Orderings for Infinite Utility Streams with Extended Anonymity*

Tapan Mitra and Kaushik Basu

6.1 Introduction

In ranking social states, which are specified by infinite utility streams, it is customary to use a social welfare quasi-ordering (SWQ), a reflexive and transitive binary relation on the social states, satisfying two widely accepted guiding principles. The equal treatment of all generations, proposed by Ramsey (1928), is formalized in the Finite Anonymity Axiom. The positive sensitivity of the social preference structure to the well-being of each generation is reflected in the Pareto Axiom.

The Finite Anonymity axiom says that if $x = (x_1, x_2, \ldots)$ and $y = (y_1, y_2, \ldots)$ are infinite utility streams, and x can be obtained by applying a finite permutation to y, then x should be declared indifferent to y. Many authors have felt that a stronger notion than the Finite Anonymity Axiom is needed to reflect intergenerational equity in intertemporal preferences.[1] This essentially means that in comparing infinite utility streams, indifference would be postulated for a larger class of permutations,[2] which would include finite permutations as a special case.

The problem with postulating indifference with respect to *arbitrary infinite permutations* is, of course, that preference relations with this feature would

* We would like to thank Shankar Sen for helpful conversations, and specifically for suggesting the characterization result which appears as Lemma 1 in the chapter. An earlier version of this chapter was presented at the IEA Roundtable Meeting in Hakone, Japan, 10–12 March 2005, and the present version has benefited from comments by Geir Asheim, Claude d'Aspremont, Kuntal Banerjee, Marc Fleurbaey and Wlodek Rabinowicz.

violate the Pareto axiom.[3] Thus, it is clear that the class of permutations, with respect to which indifference is postulated, would have to be restricted in some way if it is to be compatible with any given Paretian SWQ. However, somewhat surprisingly, there is no systematic study in the literature of the class of permutations, which are permissible, in the sense that every utility stream is pronounced to be indifferent to the corresponding permuted utility stream, according to the given Paretian SWQ.[4]

The basic question that arises then is the following: how would one specify this larger class of permutations? An approach followed in the literature has been to specify a class of infinite permutations and to argue that society *should be* indifferent between utility streams when one stream can be obtained from another by applying such an infinite permutation to it.[5]

The approach taken in this chapter is somewhat different. We wish to identify the class of permutations that *can be* allowed, given the very structure of the problem. That is, given that we seek a SWQ, which must satisfy the Pareto axiom, we wish to analyze the restrictions (if any) on the class of permutations with respect to which utility streams can be pronounced to be indifferent. What is involved here is a logical consistency check rather than any ethical principle.

Our analysis reveals two clear-cut restrictions. Given a Paretian SWQ, and denoting the set of *permissible permutations* associated with it by Π, we see that the Pareto axiom implies that the permutations in Π must be *cyclic*. Further, the transitivity property of the SWQ implies that the set Π (together with the operation of matrix mutliplication of infinite permutation matrices) must be a *group*.[6]

These are significant restrictions, dictated entirely by the mathematical structure of the problem. They also exhaust *all* the restrictions imposed by the nature of the problem. That is, given any group $\mathbb{Q}$ of cyclic permutations, there is a Paretian SWQ, such that the class Π of permissible permutations associated with it coincides exactly with $\mathbb{Q}$. Thus, we provide a complete characterization of permissible permutations that are consistent with the existence of a Paretian social welfare quasi-ordering on infinite utility streams.[7]

As the proof of our (sufficiency) result shows, a social welfare quasi-ordering that suffices for this purpose is exactly of the type known as the *Suppes-Sen grading principle*, except that it is defined with respect to all the permutations in the specified group, instead of the class of finite permutations.[8] Thus, the social welfare quasi-orderings we propose can be viewed as extended Suppes-Sen grading principles.

In view of our characterization result, we re-examine a notion of extended anonymity in which the rearrangements of utility streams allowed in any pair-wise comparison are *fixed-step permutations*, discussed in Fleurbaey and Michel (2003). That is, the rearrangements are restricted to a sequence of permutations within blocks of time of *equal length*. These blocks of time might be considered to be extended 'time periods' and permutations within each block might be treated just like rearrangements in finite societies. We show that this class of permutations is a group of cyclic permutations (and hence consistent with the existence of a Paretian social welfare quasi-ordering), which constitutes a strict extension of the class of finite permutations.

6.2 Preliminaries

Notation

Let $\mathbb{N}$ denote, as usual, the set of natural numbers $\{1, 2, 3, \ldots\}$, and let $\mathbb{R}$ denote the set of real numbers. Let Y denote the closed interval $[0, 1]$, and let the set $Y^{\mathbb{N}}$ be denoted by X. Then X is the domain of utility sequences that we are interested in. Hence, $x \equiv (x_1, x_2, \ldots) \in X$ if and only if $x_n \in [0, 1]$ for all $n \in \mathbb{N}$.

For $y, z \in \mathbb{R}^{\mathbb{N}}$, we write $y \geq z$ if $y_i \geq z_i$ for all $i \in \mathbb{N}$; and, we write $y > z$ if $y \geq z$, and $y \neq z$.

Definitions

A *social welfare quasi-ordering* (SWQ) is a binary relation, $\succcurlyeq$, on X, which is reflexive and transitive. We associate with $\succcurlyeq$ its symmetric and asymmetric components in the usual way. Thus, we write $x \sim y$ when $x \succcurlyeq y$ and $y \succcurlyeq x$ both hold; and, we write $x \succ y$ when $x \succcurlyeq y$ holds, but $y \succcurlyeq x$ does not hold.

A SWQ $\succcurlyeq_A$ is a *subrelation* to a SWQ $\succcurlyeq_B$ if (a) $x, y \in X$ and $x \succcurlyeq_A y$ implies $x \succcurlyeq_B y$; and (b) $x, y \in X$ and $x \succ_A y$ implies $x \succ_B y$.

Permutations

A *permutation* π is a one-to-one map from $\mathbb{N}$ onto $\mathbb{N}$. Any $x \in X$ can be viewed as a map from $\mathbb{N}$ to Y, associating with each $n \in \mathbb{N}$ the element $x_n \in Y$. The composite map $x \circ \pi$ is then a map from $\mathbb{N}$ to Y, associating with each $n \in \mathbb{N}$ an element $\pi(n)$ through the map π, and then associating the element $x_{\pi(n)} \in Y$ through the map x. Thus, if x is written as the sequence $(x_1, x_2, \ldots) \in X$, then $x \circ \pi$ is written as the sequence $(x_{\pi(1)}, x_{\pi(2)}, \ldots) \in X$.

Any permutation π can be *represented* by a permutation matrix. A *permutation matrix* $P = (p_{ij})_{i \in \mathbb{N}, j \in \mathbb{N}}$, is defined as follows:

$$(i)\ \textit{For each } i \in \mathbb{N},\ \textit{there is } j(i) \in \mathbb{N},\ \textit{such that } p_{ij(i)} = 1$$
$$\textit{and } p_{ij} = 0 \textit{ for all } j \neq j(i)$$
$$(ii)\ \textit{For each } j \in \mathbb{N},\ \textit{there is } i(j) \in \mathbb{N},\ \textit{such that } p_{i(j)j} = 1$$
$$\textit{and } p_{ij} = 0 \textit{ for all } i \neq i(j)$$

Given any permutation π, there is a permutation matrix, P, such that if $x \in X$, then $x \circ \pi = (x_{\pi(1)}, x_{\pi(2)}, \ldots)$ can be written as Px in the usual sense of matrix multiplication. Notice that for any permutation matrix P and any $x \in X$, the matrix multiplication is well-defined, since each row of P has one non-zero entry. Conversely, given any permutation matrix P, there is a permutation π defined by $\pi = Pa$, where $a = (1, 2, 3, \dots)$. We denote the set of all permutation matrices by $\wp$.

A *finite permutation* π is a permutation, such that there is some $N \in \mathbb{N}$, with $\pi(n) = n$ for all $n > N$. The set of all finite permutations[9] is denoted by $\mathbb{F}$.

It is useful to recall some basic properties of permutation matrices.[10] (i) If $P, Q \in \wp$, then $PQ \in \wp$. (ii) The infinite identity matrix, I, belongs to $\wp$, and for each $P \in \wp$, we have $PI = IP = P$. (iii) Given any $P \in \wp$, the transpose of P, denoted by P', belongs to P, and $PP' = P'P = I$, so that P' is the inverse of P. (iv) Finally, for $P, Q, R \in \wp$, we have:

$$P(QR) = (PQ)R$$

Thus, $\wp$ is a *group* under the usual matrix multiplication operation.[11]

The n-th unit vector in X is the sequence in X with 1 in the n-th place and 0 elsewhere, and is denoted by e^n for each $n \in \mathbb{N}$. The set of unit vectors $\{e^1, e^2, \dots\}$ is denoted by U.

If $x \in X$, then x can be written as:

$$x = \sum_{n=1}^{\infty} x_n e^n$$

where the infinite sum is interpreted as the co-ordinate-wise convergence limit of the finite sum $\sum_{n=1}^{N} x_n e^n$ as $N \to \infty$.

If $P \in P$, and $x \in X$, then (in view of the above representation of x) properties of the rearranged sequence Px can be studied by seeing how the permutation P acts on the unit vectors of X. Note that given any $e^n \in U$, the permutation P transforms the unit vector e^n to a unit vector (possibly different from e^n). Thus, the permutation matrix P maps U to U. We can, therefore, consider repeated applications of P to U, and these iterates would also remain in U. Given any $n \in N$, we can consider the sequence:

$$(Pe^n, P^2e^n, \dots)$$

generated by iterates of P applied to the unit vector e^n. The sequence is called *non-wandering* if there exist $i, j \in \mathbb{N}$, with $i < j$, such that $P^i e^n = P^j e^n$. Otherwise, it is called *wandering*. As the name suggests, a wandering sequence never revisits a point.

Non-wandering sequences can be characterized more simply as follows. Denoting $(j - i)$ by k, we see that by applying P' repeatedly (i times) to the equation $P^i e^n = P^j e^n$, we would get $e^n = P^k e^n$. Thus, a non-wandering sequence returns to e^n after a finite number of iterations. Its structure therefore is of the form of an infinitely repeated *cycle* $(Pe^n, P^2e^n, \dots, P^k e^n (= e^n), Pe^n, P^2e^n, \dots, e^n, \dots)$. If m is the smallest integer for which $P^m e^n = e^n$, then m is called the *period* of the cycle.

If P is a permutation such that for each unit vector $e^n \in U$, its iterates generate a non-wandering sequence, then P is called *cyclic*. Thus, a cyclic permutation P generates an infinitely repeated cycle, starting with every unit vector. (Notice that, in general, the period of the cycle generated might be different for different unit vectors.)

A useful property of a cyclic permutation P is that its inverse is also cyclic. To see this, consider an arbitrary unit vector $e^k \in U$. Since P is cyclic, there is some $m \in N$ such that $P^m e^k = e^k$. Applying $Q = P' = P^{-1}$ to this equation repeatedly (m times), we get $e^k = Q^m e^k$. Thus, Q is cyclic.

It is easy to check that any *finite permutation* is cyclic. On the other hand, there exist infinite permutation matrices, which are not cyclic. That is, the class $\mathbb{C}$ of cyclic permutations is a strict subset of $\wp$.

Here is a simple example of a non-cyclic permutation matrix. Let π be the permutation which maps $\mathbb{N}$ onto $\mathbb{N}$ as follows:

$$\left.\begin{array}{ll} \pi(n) = n + 2 & \text{for } n \text{ even} \\ \pi(n) = n - 2 & \text{for } n > 1 \text{ and odd} \\ \pi(1) = 2 & \end{array}\right\}$$

Note that if P is the permutation matrix associated with π then the iterates of P, when applied to the first unit vector, e^1, will generate the sequence $(e^2, e^4, e^6, \ldots)$, clearly a wandering sequence. Thus, P is not cyclic.

6.3 On Paretian SWQs with extended anonymity: necessary conditions

Given a social welfare quasi-ordering $\succcurlyeq$ on X, the set of its *permissible permutations* is defined to be:

$$\Pi(\succcurlyeq) = \{P \in \wp : Px \sim x \quad \text{for all } x \in X\}$$

That is, it is the class of permutations with respect to which every utility stream is pronounced to be indifferent to the corresponding permuted utility stream. Note that since the infinite identity matrix, I, belongs to $\wp$, and $\succcurlyeq$ is reflexive, I belongs to $\Pi(\succcurlyeq)$, so that $\Pi(\succcurlyeq)$ is always non-empty.

The standard anonymity axiom may be stated as follows.

Axiom 1. (Finite Anonymity) If $x, y \in X$, and there exist i, j in $\mathbb{N}$, such that $x_i = y_j$ and $x_j = y_i$, while $x_k = y_k$ for all $k \in \mathbb{N}$, such that $k \neq i, j$, then $x \sim y$.

It is easy to see that a SWQ $\succcurlyeq$ satisfies the Finite Anonymity axiom if and only if for every *finite* permutation $P \in \mathbb{F}$, and every $x \in X$, we have $Px \sim x$. That is, $\succcurlyeq$ satisfies the Finite Anonymity axiom if $\mathbb{F} \subset \Pi(\succcurlyeq)$.

The standard anonymity axiom suggests that we can write an extended anonymity axiom in the following way, with respect to a class $\mathbb{Q} \subset \wp$ of permutations, where $\mathbb{F} \subset \mathbb{Q}$.

Axiom 2. ($\mathbb{Q}$ – Anonymity) If $\mathbb{F} \subset \mathbb{Q} \subset \wp$, then for every $x \in X$, we have $Px \sim x$ if $P \in \mathbb{Q}$.

That is, $\succcurlyeq$ satisfies Q-Anonymity (where $\mathbb{F} \subset \mathbb{Q} \subset \wp$) if $Q \subset \Pi(\succcurlyeq)$.

We are interested in SWQs on X, which satisfy the well-known Pareto Axiom.

Axiom 3. (Pareto) If $x, y \in X$, and there is some $j \in \mathbb{N}$, such that $x_j > y_j$, while $x_k \geq y_k$ for all $k \neq j$, then $x \succ y$.

SWQs, satisfying the Pareto axiom, are called *Paretian* SWQs.

Permissible extensions of finite anonymity: two results

The question we seek to address in this subsection is the following. Given a Paretian SWQ $\succcurlyeq$, what properties are satisfied by the set of its *permissible permutations*, $\Pi(\succcurlyeq)$? Unlike the literature, we do not *postulate* any form of anonymity axiom, but rather seek to *identify* the class of permutations under which the given relation pronounces every utility stream to be indifferent to the corresponding permuted utility stream.

We obtain two restrictions that Π must satisfy.[12] First, every $P \in \Pi$ must be cyclic; second the set Π (together with the usual operation of matrix multiplication) must constitute a group.[13] We take up each of these results in turn.

For the first result, we provide, in fact, a complete characterization of cyclic permutations, which might be of independent interest.

Lemma 1 *A permutation $P \in \wp$ is cyclic if and only if there is no $x \in X$ satisfying $Px > x$.*

Proof. Suppose $P \in \wp$ is cyclic, but there is some $x \in X$ satisfying $Px > x$. Then we can find a unit vector e^k and a positive real number ε, such that:

$$Px - x \geq \varepsilon e^k \tag{6.1}$$

This yields the sequence of inequalities:

$$\begin{aligned} Px - x &\geq \varepsilon e^k \\ P^2x - Px &\geq \varepsilon P e^k \\ P^3x - P^2x &\geq \varepsilon P^2 e^k \\ &\cdots \end{aligned} \tag{6.2}$$

Let m be the period of the cycle of P. Summing the inequalities in (6.2) for $N = sm$, where $s \in \mathbb{N}$,

$$P^N x - x = \sum_{n=1}^{N} [P^n x - p^{n-1} x] \geq \varepsilon s \left[\sum_{n=1}^{m} P^n e^k \right] \tag{6.3}$$

Denoting the sequence $(1, 1, 1, \ldots)$ by e, we have from (6.3),

$$(e/s) \geq \varepsilon \left[\sum_{n=1}^{m} P^n e^k \right] \quad \textit{for all } s \in \mathbb{N} \tag{6.4}$$

But the vector on the left-hand side of (6.4) goes to zero as $s \to \infty$, while the right-hand side of (6.4) is a non-negative non-zero vector independent of s. This contradiction establishes the necessity part of the result.

To establish sufficiency, suppose that $P \in \wp$ is not cyclic. Then, denoting the inverse of P by Q, we know that Q cannot be cyclic. Thus, there is some unit vector $e^k \in U$, for which the sequence $(Qe^k, Q^2 e^k, \ldots)$ is wandering. Each vector in this sequence is a unit vector. Since the sequence is wandering, any unit vector occurs at most once in the sequence. Thus, the sequence $(x(1), x(2), \ldots)$ defined by:

$$x(N) = \sum_{n=1}^{N} Q^n e^k \quad \textit{for } N \in \mathbb{N} \tag{6.5}$$

is a monotonic non-decreasing sequence in X, bounded above by $e = (1, 1, 1, \ldots)$. Consequently, $x(N)$ has a (co-ordinate-wise convergence) limit as $N \to \infty$. Define this limit by x; then $x \in X$.

Multiplying through in (6.5) by Q, we have:

$$Qx(N) = \sum_{n=1}^{N} Q^{n+1} e^k \quad \textit{for } N \in \mathbb{N} \tag{6.6}$$

Subtracting (6.6) from (6.5), for each $N \in \mathbb{N}$,

$$x(N) - Qx(N) = Qe^k - Q^{N+1} e^k \tag{6.7}$$

Taking co-ordinate-wise convergence limits in (6.7), we obtain:

$$x - Qx = Qe^k \tag{6.8}$$

Multiplying through in (6.8) by P, we get:

$$Px - x = e^k > 0$$

This completes the sufficiency part of the proof.

We now note a principal implication of this characterization of cyclic permutations.

Proposition 1 *Suppose $\succcurlyeq$ is a Paretian SWQ. Then, every $P \in \Pi(\succcurlyeq)$ must be cyclic.*

Proof. Given $\succcurlyeq$, denote $\Pi(\succcurlyeq)$ by Π. Suppose, contrary to the proposition, there is some $P \in \Pi$, which is not cyclic. Then, by Lemma 1, there is $x \in X$ such that $Px > x$. Since $P \in \Pi$, we must have $Px \sim x$. But, since $Px \in X$ and $Px > x$, we must have $Px \succ x$ because $\succcurlyeq$ is Paretian. This contradiction establishes the result.

The second result, while fairly straightforward to establish, provides a restriction, which is more involved and therefore harder to check.

Proposition 2 *Suppose $\succcurlyeq$ is a Paretian SWQ. Then, $\Pi(\succcurlyeq)$ is a group with respect to the operation of matrix multiplication.*

Proof. Given $\succcurlyeq$, denote $\Pi(\succcurlyeq)$ by Π. We check the four properties which define a group. First, let P, Q belong to Π. Define $R = PQ$; we know that $R \in \wp$. We have to show that $R \in \Pi$. Let $x \in X$ be arbitrarily specified. Then, since $Q \in \Pi$, we have $Qx \sim x$. Denoting Qx by y, we note that $y \in X$, and since $P \in \Pi$, we also have $Py \sim y$. Denoting Py by z, we note that $z \in X$, and $z \sim y$ while $y \sim x$, so that $z \sim x$ since $\succcurlyeq$ is transitive. Thus, $PQx = Py = z$ is indifferent to x. So, $R = PQ$ must belong to Π.

Second, the identity matrix $I \in \Pi$ (by definition of $\Pi(\succcurlyeq)$, since $\succcurlyeq$ is reflexive) and given any $P \in \Pi$, we have $PI = IP = P$, since $P \in \wp$.

Third, if $P \in \Pi$, then $P' \in \wp$, and we have to show that $P' \in \Pi$. Let x be an arbitrary point in X. Then, defining $y = P'x$, we see that $y \in X$. Further, multiplying both sides of this equation by P, we see that $Py = PP'x = x$ (since P' is the inverse of P). Since $P \in \Pi$, we must have $Py \sim y$; this means that $x \sim P'x$. Since $x \in X$ was arbitrary, this shows that $P' \in \Pi$.

Finally, if $P, Q, R \in \Pi$, then $P, Q, R \in \wp$, and so $(PQ)R = P(QR)$.

Permissible extensions of finite anonymity: two examples

The restrictions imposed by the above propositions on the set of permissible permutations of Paretian SWQs are significant ones. We illustrate this point by discussing two examples. Consider, first, the example, introduced in section 6.2.3.

Example 1: Let π be the permutation which maps $\mathbb{N}$ onto $\mathbb{N}$ as follows:

$$\left.\begin{array}{l} \pi(n) = n + 2 \ \ \textit{for } n \textit{ even} \\ \pi(n) = n - 2 \ \ \textit{for } n > 1 \textit{ and odd} \\ \pi(1) = 2 \end{array}\right\} \qquad (6.9)$$

We have already noted that if P is the permutation matrix associated with π then P is not cyclic. Consequently P cannot belong to $\Pi(\succcurlyeq)$ if $\succcurlyeq$ is any Paretian SWQ.

Perhaps a more transparent way to look at the permutation defined above is to see the effect of it on a particular utility sequence $x \in X$:

$$x = (0, 1, 0, 1, 0, 1, 0, \dots)$$
$$Px = (1, 1, 0, 1, 0, 1, 0, \dots)$$

Clearly, what the permutation effectively does is to produce a Pareto superior utility sequence.

Proposition 1 shows that for every Paretian SWQ, $\Pi(\succcurlyeq)$ must be a subset of the class $\mathbb{C}$ of cyclic permutations. Thus, for every Paretian SWQ, $\Pi(\succcurlyeq)$ must be a strict subset of $\wp$, since the class $\mathbb{C}$ of cyclic permutations is a strict subset of $\wp$.

One might wonder whether it is possible to have a Paretian SWQ $\succcurlyeq$, for which $\Pi(\succcurlyeq)$ is $\mathbb{C}$. Unfortunately, $\mathbb{C}$ is not a group, as the following example shows. Thus, for every Paretian SWQ $\succcurlyeq$, the set of permissible permutations $\Pi(\succcurlyeq)$ must exclude some cyclic permutation, and $\Pi(\succcurlyeq)$ must be a strict subset of $\mathbb{C}$.

Example 2: Let π_1 be a permutation, defined as follows:

$$\left.\begin{array}{l} \pi_1(n) = n + 1 \;\; \textit{if } n \textit{ is odd} \\ \pi_1(n) = n - 1 \;\; \textit{if } n \textit{ is even} \end{array}\right\} \tag{6.10}$$

Clearly the permutation matrix P_1 associated with π_1 is cyclic, with a cycle of period 2 for each unit vector.

Let π_2 be the permutation, defined as follows:

$$\left.\begin{array}{l} \pi_2(1) = 1 \\ \pi_2(n) = n + 1 \;\; \textit{if } n \textit{ is even} \\ \pi_2(n) = n - 1 \;\; \textit{if } n > 1 \textit{ and odd} \end{array}\right\} \tag{6.11}$$

Clearly, the permutation matrix P_2 associated with π_2 is cyclic, with a cycle of period 2 for each unit vector, starting with the second one; it has a cycle of period 1 for the first unit vector.

While P_1 and P_2 belong to the set $\mathbb{C}$ of cyclic permutations, it is easy to check that the composite permutation $\pi_2 \circ \pi_1$ is precisely the permutation π of Example 1, so that $P_2 P_1 = P$ is not cyclic.

Again, it is instructive to look at the effect of these permutations on a specific utility sequence $x \in X$:

$$x = (0, 1, 0, 1, 0, 1, 0, \dots)$$
$$P_1 x = (1, 0, 1, 0, 1, 0, 1, \dots)$$
$$P_2 P_1 x = (1, 1, 0, 1, 0, 1, 0, \dots)$$

If P_1 and P_2 both belong to $\Pi(\succcurlyeq)$, for some Paretian SWQ $\succcurlyeq$ then $P_1 x \sim x$ and $P_2(P_1 x) \sim (P_1 x)$, and one might not find either of these binary comparisons to be unacceptable. However, since the SWQ $\succcurlyeq$ is transitive, we must then have $P_2(P_1 x) \sim x$, and this is clearly unacceptable since $P_2(P_1 x)$ is Pareto superior to x.

6.4 On Paretian SWQs with extended anonymity: sufficient conditions

We have noted above that the very structure of our problem imposes significant restrictions on the class of permissible permutations associated with any Paretian SWQ. Now, we ask whether the restrictions obtained in Propositions 1 and 2 exhaust all the restrictions on the class of permissible permutations associated with any Paretian SWQ. In other words, if $\mathbb{Q}$ is an *arbitrary* group of cyclic permutations, can we always define a Paretian SWQ $\succcurlyeq$ for which the class of permissible permutations $\Pi(\succcurlyeq)$ coincides exactly with $\mathbb{Q}$? The answer to this question is (somewhat surprisingly) in the affirmative, so that we have, in fact, a complete characterization of the class of permissible permutations associated with any Paretian SWQ.

Our demonstration of the above result consists in writing down a binary relation $\succcurlyeq$ and checking that (i) it is a Paretian SWQ, and that (ii) $\Pi(\succcurlyeq) = \mathbb{Q}$. The particular binary relation we use is exactly of the form of the Suppes-Sen grading principle, but with the set of finite permutations replaced by the given group of cyclic permutations, $\mathbb{Q}$.

Formally, for every specification of a group of cyclic permutations $\mathbb{Q}$, we have a corresponding '$\mathbb{Q}$-*grading principle*', denoted by $\succcurlyeq_{\mathbb{Q}}$, defined as follows: if $x, y \in X$, then $x \succcurlyeq_{\mathbb{Q}} y$ if and only if there is some $P \in \mathbb{Q}$ such that $Px \geq y$. The symmetric ($\sim_{\mathbb{Q}}$) and asymmetric ($\succ_{\mathbb{Q}}$) parts of $\succcurlyeq_{\mathbb{Q}}$ are defined in the usual way.

Proposition 3 *Let $\mathbb{Q}$ be a group of cyclic permutations. Then, the $\mathbb{Q}$-grading principle is a Paretian social welfare quasi-ordering satisfying $\mathbb{Q}$-Anonymity (that is, $\Pi(\succcurlyeq_{\mathbb{Q}}) \subset \mathbb{Q}$). Moreover, $\mathbb{Q} \subset \Pi(\succcurlyeq_{\mathbb{Q}})$.*

Proof. We check first that the binary relation $\succcurlyeq_{\mathbb{Q}}$ is reflexive and transitive, so that it constitutes a social welfare quasi-ordering.

Let $x \in X$. Since the identity matrix $I \in \mathbb{Q}$, and $Ix \geq x$, we have $x \succsim_{\mathbb{Q}} x$, verifying that $\succsim_{\mathbb{Q}}$ is reflexive.

Let $x, y, z \in X$ with $x \succsim_{\mathbb{Q}} y$ and $y \succsim_{\mathbb{Q}} z$. Then, there exist $P \in \mathbb{Q}$ and $Q \in \mathbb{Q}$ such that $Px \geq y$ and $Qy \geq z$. Since $P, Q \in \mathbb{Q}$ and $\mathbb{Q}$ is a group, $R \equiv QP \in \mathbb{Q}$. Applying the permutation Q to the inequality $Px \geq y$, we get $QPx \geq Qy$, and using the inequality $Qy \geq z$, we get $QPx \geq z$. Thus, we have $R \in \mathbb{Q}$ and $Rx \geq z$, so that $x \succsim_{\mathbb{Q}} z$, establishing transitivity of $\succsim_{\mathbb{Q}}$.

We now show that $\succsim_{\mathbb{Q}}$ is Paretian. Let $x, y \in X$ with $x > y$. Then since the identity matrix $I \in \mathbb{Q}$, and $Ix = x > y$, we certainly have $x \succsim_{\mathbb{Q}} y$. We claim now that $y \succsim_{\mathbb{Q}} x$ does not hold. For, if $y \succsim_{\mathbb{Q}} x$, then there is some $P \in \mathbb{Q}$ such that $Py \geq x$. But, since $x > y$, we must then have $Py > y$. By Lemma 1, this contradicts the fact that P is cyclic. Thus, $x \succsim_{\mathbb{Q}} y$ holds and $y \succsim_{\mathbb{Q}} x$ does not hold, and so $x \succ_{\mathbb{Q}} y$.

Next, we show that $\mathbb{Q}$ is a subset of $\Pi(\succsim_{\mathbb{Q}})$, the set of permissible permutations associated with $\succsim_{\mathbb{Q}}$. Let $P \in \mathbb{Q}$, and let x be an arbitrary point in X. Define $y \equiv Px$. Since $Px = y$, we clearly have $x \succsim_{\mathbb{Q}} y$. Also, multiplying the equation $Px = y$ by P', we have $x = P'Px = P'y$. Since $\mathbb{Q}$ is a group, $P' \in \mathbb{Q}$, so we must have $y \succsim_{\mathbb{Q}} x$. Thus, $y \sim_{\mathbb{Q}} x$; that is, $Px \sim_{\mathbb{Q}} x$.

Finally, we show that $\Pi(\succsim_{\mathbb{Q}})$ is a subset of $\mathbb{Q}$. Let $P \in \wp$ and suppose that $Px \sim_{\mathbb{Q}} x$ for all $x \in X$. Choose $\overline{x} = (\overline{x}_n)$, where $\overline{x}_n = 1/2^{n-1}$ for all $n \in \mathbb{N}$. Clearly $\overline{x} \in X$, and $P\overline{x} \sim_{\mathbb{Q}} \overline{x}$. Define $\overline{y} = P\overline{x}$; then $\overline{y} \in X$. Since $\overline{y} \sim_{\mathbb{Q}} \overline{x}$, there is $Q \in \mathbb{Q}$ and $R \in \mathbb{Q}$ such that $Q\overline{x} \geq \overline{y}$ and $R\overline{y} \geq \overline{x}$. Multiplying the latter inequality by $R' \in Q$, we have $\overline{y} \geq R'\overline{x}$. Summarizing, we have:

$$z \equiv Q\overline{x} \geq \overline{y} \geq R'\overline{x} \equiv z' \tag{6.12}$$

We can write:

$$z_n = z'_n + (z_n - z'_n) \ \textit{for all } n \in \mathbb{N} \tag{6.13}$$

and sum (6.13) from $n = 1$ to $n = N$, where $N \in \mathbb{N}$, to obtain:

$$\sum_{n=1}^{N} Z_n = \sum_{n=1}^{N} Z'_n + \sum_{n=1}^{N} (Z_n - Z'_n) \tag{6.14}$$

Note that z and z' are rearrangements of the sequence $\overline{x}$ and since $\sum_{n=1}^{N} \overline{x}$, is absolutely convergent (as $N \to \infty$) with a sum equal to 1, $\sum_{n=1}^{N} z^n$ and $\sum_{n=1}^{N} z'_n$ must both converge to 1 as $N \to \infty$.[14] Using (6.12), $\sum_{n=1}^{N} (z_n - z'_n)$ is a monotonically non-decreasing sequence (in N) bounded above by 1, and must converge. Taking limits in (6.14), we must have $\sum_{n=1}^{N} (z_n - z'_n)$ converging to zero as $N \to \infty$. But, since $(z_n - z'_n) \geq 0$ for each $n \in \mathbb{N}$, this is only possible if $(z_n - z'_n) = 0$ for each $n \in \mathbb{N}$. Thus, $z = z'$, and so by (6.12) we have:

$$z \equiv Q\overline{x} = \overline{y} = R'\overline{x} = z' \tag{6.15}$$

In particular, we get $\bar{y} = Q\bar{x}$ from (6.15). But by definition $\bar{y} = P\bar{x}$. Thus, we must have:

$$P\bar{x} = Q\bar{x} \tag{6.16}$$

Since $\bar{x}_i \neq \bar{x}_j$ whenever $i, j \in \mathbb{N}$ with $i \neq j$, (6.16) can hold only if $Q = P$. Thus, $P \in \mathbb{Q}$, finishing the proof of the Proposition.

Remark: Proposition 3 establishes that, given a group of cyclic permutations $\mathbb{Q}$, the $\mathbb{Q}$-grading principle $\succcurlyeq_{\mathbb{Q}}$, is a Paretian SWQ, which satisfies $\mathbb{Q}$-Anonymity. It can be used to show that a SWQ $\succcurlyeq$ satisfies the Pareto axiom and the $\mathbb{Q}$-Anonymity axiom if and only if $\succcurlyeq_{\mathbb{Q}}$ is a subrelation to $\succcurlyeq$. That is, the $\mathbb{Q}$-grading principle is the least restrictive SWQ satisfying the Pareto axiom and the $\mathbb{Q}$-Anonymity axiom. This observation has been noted by Banerjee (2006).

6.5 On a group of cyclic permutations

Our characterization of possible extensions of anonymity, consistent with a Paretian SWQ, has not addressed one central question. Is there a group of cyclic permutations which is a strict extension of the class of finite permutations? In this section, we address this question by specifying a group of cyclic permutations which has several attractive properties. First, it includes the class of finite permutations. Second, it strictly extends the class of finite permutations by allowing infinite permutations, which can essentially be written as a sequence of finite permutations over blocks of time of equal length. Third, it includes the class of infinite permutations that has most commonly been proposed in extensions of the standard anonymity axiom.

Our class of permutations has to be carefully chosen in view of the restrictions imposed by Propositions 1 and 2. While the restriction of being cyclic is relatively easy to check, the restriction of being a group is more subtle, since it pertains to compositions of permutations. This difference between the two (independent) restrictions is most clearly displayed in Examples 1 and 2. Note that Example 2 shows that even if we choose the class of permutations $\mathbb{Q}$ to be the subset of $\mathbb{C}$, consisting only of cyclic permutations with the period of cycles uniformly bounded above (independent of the unit vector chosen), it would not satisfy the second restriction.

We now proceed to define formally our class of permutations. Following Fleurbaey and Michel (2003), we call them *fixed step permutations*. Given a permutation matrix, $P \in \wp$, and $n \in \mathbb{N}$, we denote the $n \times n$ matrix $(p_{ij})_{i,j \in \{1,\ldots,n\}}$ by $P(n)$. Let $\mathbb{S} = \{P \in \wp$: there is some $k \in \mathbb{N}$, such that for each $n \in \mathbb{N}$, $P(nk)$ is a finite dimensional permutation matrix$\}$.

If $P, Q \in \mathbb{S}$, then there are $k \in \mathbb{N}, k' \in \mathbb{N}$, such that for each $n \in \mathbb{N}$, $P(nk)$ and $Q(nk')$ are finite dimensional permutation matrices. Define $R = PQ$. Then $R \in \wp$. Further, defining $k'' = kk'$, we can check that for each $n \in N$, $R(nk'')$

is a finite dimensional permutation matrix. Thus, $R \in S$. Now, it is easy to check that $\mathbb{S}$ is also a group.

If $P \in \mathbb{S}$, then P is clearly cyclic since the iterates of P acting on any unit vector will return to the unit vector in at most k iterations. Thus, $\mathbb{S}$ is a group of cyclic permutations.

If P represents a permutation in $\mathbb{F}$, then there is some $k \in \mathbb{N}$ such that $P(k)$ is a finite dimensional permutation matrix and $p_{ii} = 1$ for all $i > k$. Thus, we certainly have $P(nk)$ to be a finite dimensional permutation matrix for each $n \in \mathbb{N}$. Thus, $\mathbb{S}$ includes the class $\mathbb{F}$ of finite permutations.

One of the most common examples considered in proposing an extension of the Anonymity axiom is the following:

$$x = (0, 1, 0, 1, 0, 1, 0, 1, \ldots)$$
$$y = (1, 0, 1, 0, 1, 0, 1, 0, \ldots)$$

Although x cannot be obtained from y (nor y from x) by applying a finite permutation, it has been felt that x should be declared indifferent to y. That is, at least this class of (infinite) permutation should be allowed in any extended notion of Anonymity. We see that for the (infinite) permutation P involved here, $P(2n)$ is a finite dimensional permutation matrix for each $n \in \mathbb{N}$, and so P belongs to $\mathbb{S}$.

We do not know whether the group of fixed step permutations, $\mathbb{S}$, is a *maximal* group of cyclic permutations. In fact, it would be useful to know whether there are other groups of cyclic permutations, which have all the three properties stated above. If not, there is a strong case for focusing exclusively on the group of fixed step permutations, in discussions of extended anonymity.

Notes

1. See, for example, Lauwers (1995, 1997), Liedekerke and Lauwers (1997), Fleurbaey and Michel (2003).
2. In what follows, we use the terms 'permutations' and 'permutation matrices' interchangeably. The connection between the two is the following. A permutation is a one-to-one map from the natural numbers onto the natural numbers. Any such permutation can be represented by a permutation matrix. See section 6.2 for a discussion.
3. This point is well recognized in the literature. See, for example, Lauwers (1997), and Asheim and Tungodden (2004).
4. Fleurbaey and Michel (2003) have undertaken a very comprehensive study of anonymity with respect to infinite permutations. However, their approach is to specify a class of permutations (they consider fixed step, variable step and finite length permutations) and ask whether indifference with respect to this class is consistent with axioms like Pareto or Weak Pareto or Continuity. Our approach

treats the class of permutations as a 'choice variable' and seeks to characterize the class which is compatible with Paretian SWQs.

5. See the papers by Lauwers (1995, 1998), where he considers a class of permutations π (which he calls 'bounded permutations') which satisfy $(\pi(n)/n) \to 1$ as $n \to \infty$. It is not quite clear, though, why this class is of special interest from the point of view of intergenerational equity.
6. The terms 'permissible permutations', 'cyclic permutations' and 'group of permutation matrices' are formally defined in Section 6.2.
7. The framework in which our result is established is, by now, the standard one, employed, for instance, in Diamond (1965), Svensson (1980) and Basu and Mitra (2003).
8. The grading principle is due to Suppes (1966). For a comprehensive analysis of it, see Sen (1970). Svensson (1980) provides a formal definition of the Suppes-Sen grading principle in the context of infinite utility streams. It can be characterized as the least restrictive SWQ satisfying the Pareto and Anonymity axioms; see d'Aspremont (1985) and Asheim, Buchholz and Tungodden (2001).
9. For properties of finite permutations, see, for example, Hohn (1973).
10. Some of the basic properties of infinite permutation matrices can be found in Cooke (1950).
11. A *group* is a set of objects, $\Im$, together with a binary operation $\otimes$ on $\Im$ such that:
    (i) If $A, B \in \Im$, then $A \otimes B \in \Im$.
    (ii) An identity element, $I \in \Im$, such that for every $A \in \Im$, $I \otimes A = A, \otimes I = A$.
    (iii) For every $A \in \Im$, there is $A' \in \Im$, such that $A \otimes A' = A' \otimes A = I$.
    (iv) If $A, B, C \in \Im$, then $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.
12. When there is no danger of confusion, we will denote $\Pi(\succsim)$ by Π, it being understood that Π is associated with the SWQ $\succsim$ given in the relevant context. This simplifies the notation.
13. Infinite permutation matrices have not been systematically studied in the mathematics literature, which focuses almost exclusively on one problem: what is the class of rearrangements which will preserve the sum of a conditionally convergent series? See Schaefer (1981) and the references cited in his paper. This problem arose from a famous result of Riemann that a rearrangement of a conditionally convergent series can be convergent to any pre-specified number or even divergent.
14. See, for example, Rudin (1976, p. 78).

References

Asheim, G.B., W. Buchholz and B. Tungodden (2001) 'Justifying Sustainability', *Journal of Environmental Economics and Management*, vol. 41, pp. 252–68.

Asheim, G.B. and B. Tungodden (2004) 'Resolving Distributional Conflicts between Generations', *Economic Theory*, vol. 24, pp. 221–30.

d'Aspremont, C. (1985) 'Axioms for Social Welfare Orderings', in L. Hurwicz, D. Schmeidler and H. Sonnenschein (eds), *Social Goals and Social Organizations: Essays in Memory of Elisha Pazner* (Cambridge: Cambridge University Press), pp. 19–76.

Banerjee, K. (2006) 'On the Extension of the Utilitarian and Suppes-Sen Social Welfare Relations to Infinite Utility Streams', *Social Choice and Welfare*, vol. 27, pp. 327–39.

Basu, K. and T. Mitra (2003) 'Aggregating Infinite Utility Streams with Intergenerational Equity: The Impossibility of Being Paretian', *Econometrica*, vol. 71, pp. 1557–63.
Cooke, R.G. (1950) *Infinite Matrices and Sequence Spaces* (London: Macmillan).
Diamond, P. (1965) 'The Evaluation of Infinite Utility Streams', *Econometrica*, vol. 33, pp. 170–7.
Fleurbaey, M. and P. Michel (2003) 'Intertemporal Equity and the Extension of the Ramsey Principle', *Journal of Mathematical Economics*, vol. 39, pp. 777–802.
Hohn, F.E. (1973) *Elementary Matrix Algebra* (New York: Macmillan).
Lauwers, L. (1995) 'Time-Neutrality and Linearity', *Journal of Mathematical Economics*, vol. 24, pp. 347–51.
Lauwers, L. (1997) 'Rawlsian Equity and Generalized Utilitarianism with an Infinite Population', *Economic Theory*, vol. 9, pp. 143–50.
Lauwers, L. (1998) 'Intertemporal Objective Functions: Strong Pareto versus Anonymity', *Mathematical Social Sciences*, vol. 35, pp. 37–55.
Liedekerke, L. and L. Lauwers (1997) 'Sacrificing the Patrol: Utilitarianism, Future Generations and Infinity', *Economics and Philosophy*, vol. 13, pp. 159–74.
Ramsey, F.P. (1928) 'A Mathematical Theory of Savings', *Economic Journal*, vol. 38, pp. 543–59.
Rudin, W. (1976) *Principles of Mathematical Analysis* (New York: McGraw-Hill).
Schaefer, P. (1981) 'Sum-Preserving Rearrangements of Infinite Series', *American Mathematical Monthly*, vol. 88, pp. 33–40.
Sen, A.K. (1970) *Collective Choice and Social Welfare* (San Francisco: Holden-Day).
Suppes, P. (1966) 'Some Formal Models of Grading Principles', *Synthèse*, vol. 6, pp. 284–306.
Svensson, L.-G. (1980) 'Equity among Generations', *Econometrica*, vol. 48, pp. 1251–6.

7

Pareto Principle and Intergenerational Equity: Immediate Impatience, Universal Indifference and Impossibility*

Yongsheng Xu

7.1 Introduction

In an important contribution to the problem of aggregating infinite utility streams, Svensson (1980) shows the existence of a *social welfare relation* – a reflexive, transitive and complete binary relation over all possible infinite utility streams that accommodates the axioms of *Pareto* and *intergenerational equity*. This possibility result is in sharp contrast with the seminal contribution by Diamond (1965) in the same context of aggregating infinite utility streams, who established the non-existence of a *social welfare function* – a function which aggregates an infinite utility stream into a real number that satisfies the axioms of *Pareto*, *intergenerational equity* and *continuity* (in the suprametric). The axiom of continuity in Diamond's result is shown to be redundant by Basu and Mitra (2003a) recently: they show that, in aggregating infinite utility streams, there exists no social welfare function satisfying the axioms of Pareto and intergenerational equity.

The possibility result by Svensson suggests the compatibility of the Pareto principle and intergenerational equity for a social welfare relation, while the impossibility results by Diamond, and by Basu and Mitra suggest that the compatibility of the Pareto principle and intergenerational equity breaks

* I am greatly indebted to Prasanta K. Pattanaik and Kotaro Suzumura for many discussions on several related subjects we have had over the years. In preparing this version, I have benefited greatly from detailed comments by Kaushik Basu on an earlier version of the paper, and from questions, comments, and discussions of the participants, especially those of Geir Asheim, Walter Bossert, Marc Fleurbaey, Tapan Mitra and Tomoichi Shinotsuka, at the IEA Roundtable in Hakone.

down when a social welfare relation is replaced by a social welfare function. Though the possibility of accommodating both the Pareto principle and intergenerational equity for a social welfare relation in aggregating infinite utility streams can be obtained, it is not clear what structure such a social welfare relation may have and to what extent the possibility may be obtained. This is because, in proving his possibility result, Svensson uses a non-constructive method by making use of Szpilrajn's lemma on extending a reflexive and transitive binary relation to a reflexive, transitive and complete binary relation.

The purpose of this chapter is therefore twofold. First, we examine the scope of obtaining the possibility result for a social welfare relation to be both Paretian and intergenerationally equitable. We show that, under a very mild restriction on social welfare relations, it is not possible to accommodate both the Pareto principle and intergenerational equity. Therefore, the scope of social welfare relations that are both Paretian and intergenerationally equitable is rather limited. Secondly, we examine, under a set of common restrictions on a social welfare relation, the respective implications of the axiom of Pareto and of the axiom of intergenerational equity. We show that the axiom of Pareto implies *immediate impatience* – an impatience for any period t over its very next period, and the axiom of intergenerational equity implies *universal indifference*: every infinite utility stream is indifferent to any other infinite utility stream. Therefore, to some extent, our results clarify the structure of a social welfare relation satisfying the Pareto principle and intergenerational equity. The organization of the chapter is as follows. Section 7.2 presents the basic notation and definitions. Section 7.3 introduces our basic axioms and presents our impossibility result. Section 7.4 examines the respective implications of the axioms of Pareto and intergenerational equity. A brief conclusion is contained in section 7.5.

7.2 Notation

$\mathbb{R}$ is to denote the set of all real numbers, and $\mathbb{N}$ is to denote the set of all positive integers. Let X be a non-empty subset of $\mathbb{R}$ containing at least two elements. The set of all infinite utility streams is to be denoted by X^∞. The elements in X^∞ are the infinite utility streams and for $x = (x_1, x_2, \ldots, x_m, \ldots) \in X^\infty$, x_m is the utility of generation $m \in \mathbb{N}$. For all $x = (x_1, x_2, \ldots, x_m, \ldots)$ and $y = (y_1, y_2, \ldots, y_m, \ldots) \in X^\infty$, $x > y$ if and only if $[x_m \geq y_m$ for all $m \in \mathbb{N}$ and $x \neq y]$.

For any $t \in \mathbb{N}$, and for any $x_1, \ldots, x_t \in X$, let ${}_1x_t = (x_1, \ldots, x_t)$. For any $t \in \mathbb{N}$, any ${}_1x_t$, and any $y \in X^\infty$, an infinite stream $({}_1x_t, y)$ will mean $(x_1, x_2, \ldots, x_t, y_1, y_2, \ldots, y_m, \ldots)$. For any $t \in \mathbb{N}$ and any $x_1, \ldots, x_t \in X$, let $({}_1x_{trep})$ denote the infinite utility stream where ${}_1x_t$ repeats infinite times. For all $i, j \in \mathbb{N}$ and all $x \in X^\infty$, $x(ij)$ is the infinite utility stream obtained from x by switching utilities of generations i and j while keeping utilities of all other generations unchanged.

Let $\succeq$ be a reflexive, transitive and complete binary relation over X^∞. The symmetric and asymmetric of $\succeq$ will be denoted by $\sim$ and $\succ$, respectively. $\succeq$ is referred to as a *social welfare relation*.

7.3 Basic axioms and an impossibility result

Basic axioms

Consider the following axioms to be imposed on a social welfare relation $\succeq$.

Pareto (P) For all $x, y \in X^\infty$, if $x > y$ then $x \succ y$.

Very Weak Dominance (VWD) For all $x, y \in X^\infty$, if $[x_1 > y_1$ and $x_i = y_i$ for all $i > 1]$, then $x \succ y$.

Intergenerational Equity (IE) For all $x \in X^\infty$ and all $i, j \in N$, $x \sim x(ij)$.

Support (S) For all $x = (x_1, \ldots, x_m, \ldots)$, $y = (y_1, \ldots, y_m, \ldots) \in X^\infty$, if $x \succ y$, then there exists $t_0 \in \mathbb{N}$ such that $[({}_1x_t, z) \succ ({}_1y_t, z)$ for all $z \in X^\infty$ and all $t \geq t_0]$.

Minimal Support (MS) For all $x = (x_1, \ldots, x_m, \ldots)$, $y = (y_1, \ldots, y_m, \ldots) \in X^\infty$, if $x \succ y$, then there exist $t \in \mathbb{N}$, and $z, z' \in X^\infty$ such that $({}_1x_t, z) \succ ({}_1y_t, z)$ and $({}_1x_{t+1}, z') \succ ({}_1y_{t+1}, z')$.

The axiom of Pareto is the standard Pareto principle used in the literature on evaluating infinite utility streams. Very weak dominance requires that if for two infinite utility streams, x and y, the first generation in x enjoys a higher utility than the first generation in y, and the respective utility levels for all other generations under x and y are the same, then the infinite utility stream x is ranked higher than the infinite utility stream y. It is clear that very weak dominance is weaker than Pareto. It turns out that very weak dominance is sufficient for our results in this chapter. It is also interesting to note that the axiom of very weak dominance is weaker than the axiom of *weak dominance* (WD), which requires that if $x, y \in X^\infty$ are such that [for some $i \in \mathbb{N}, x_i > y_i$ and $x_k = y_k$ for all $k \in \mathbb{N} \setminus \{i\}$], then $x \succ y$, used in Basu and Mitra (2007) in establishing their possibility result. The axiom of intergenerational equity is the standard intergenerational equity condition used in the literature. It is also known as finite anonymity. The axiom of support can be regarded as a basic requirement for ranking one infinite utility stream strictly higher than another infinite utility stream.

It essentially says that the role played by any far distant 'tails' of infinite utility streams x and y in ranking x strictly higher than y is minimum and can be 'neglected': any far distant 'tails' of x and y can be replaced by any common infinite utility streams without affecting their corresponding strict rankings. The axiom of minimal support is a much weaker requirement than the axiom of support. It requires that, whenever an infinite utility stream x is ranked strictly higher than another infinite utility stream y, then there exist a generation t and infinite utility streams z and z' such that the replacement of the 'tails' of both x and y with the common z after the generation t will preserve the corresponding strict rankings, and so does the replacement of the 'tails' of x and y with the common z' after the generation $t+1$. In other words, it requires that infinite programmes be consistent with at least two finite programmes. It turns out that minimal support is sufficient for the results of this chapter.

Similar conditions to (S) and (MS) have been proposed and discussed in the literature on evaluating infinite utility streams. For example, Brock (1970)'s *Axiom 3* has the same spirit as (S) and (MS) and requires that, for all $x, y \in X^\infty$, if $x \succ y$ then there exist $n_0 \in \mathbb{N}$ and $c \in X$ such that $[n \geq n_0 \Rightarrow (x_1, \ldots, x_n, c, c, \ldots) \succ (y_1, \ldots, y_n, c, c, \ldots)]$. Fleurbaey and Michel (2003) propose a condition called *Limit Ranking*, which requires that, for all $x, y \in X^\infty$, if there exists $z \in X^\infty$ such that $({}_1x_t, z) \succeq ({}_1y_t, z)$ for all $t \in \mathbb{N}$, then $x \succeq y$.

Impossibility of a social welfare relation being Paretian and intergenerationally equitable

We now turn to the first main result of this Section, which shows the incompatibility of the axioms of very weak dominance, intergenerational equity, and minimal support.

Theorem 1. There is no $\succeq$ satisfying the axioms of (VWD), (IE) and (MS).

Proof. Suppose to the contrary that there exists $\succeq$ satisfying (VWD), (IE) and (MS). Let $a, b \in X$ with $a > b$. Consider $x = ((a, b)_{rep})$ and $y = ((b, a)_{rep})$. Since $\succeq$ is complete, there are three cases to be considered: (i) $x = ((a, b)_{rep}) \succ ((b, a)_{rep}) = y$; (ii) $y = ((b, a)_{rep}) \succ ((a, b)_{rep}) = x$; and (iii) $x = ((a, b)_{rep}) \sim ((b, a)_{rep}) = y$.

Case (i): $x = ((a, b)_{rep}) \succ ((b, a)_{rep}) = y$. If $x = ((a, b)_{rep}) \succ ((b, a)_{rep}) = y$, then (MS) implies that there exist $t \in \mathbb{N}$ and $z, z' \in X^\infty$ such that, $({}_1x_t, z) \succ ({}_1y_t, z)$ and $({}_1x_{t+1}, z') \succ ({}_1y_{t+1}, z')$. Clearly, either t or $t+1$ is even. Without loss of generality, let t be even. Then, we have $(a, b, \ldots a, b, z) \succ (b, a, \ldots b, a, z)$, where a, b and b, a appear $t/2$ times in their respective infinite utility streams. Since t is finite, by the repeated use of (IE) and the transitivity of $\succeq$, we must have $(a, b, \ldots a, b, z) \sim (b, a, \ldots b, a, z)$, a contradiction.

Case (ii): $y = ((b, a)_{rep}) \succ ((a, b)_{rep}) = x$. In this case, by employing a similar argument as that for Case (i), we can derive another contradiction.

Case (iii): $x = ((a, b)_{rep}) \sim ((b, a)_{rep}) = y$. By (VWD), $(a, (b, a)_{rep}) \succ (b, (b, a)_{rep})$. Noting that $((a, b)_{rep}) \sim ((b, a)_{rep})$ and that $(a, (b, a)_{rep}) = ((a, b)_{rep})$, the transitivity of $\succeq$ implies that $((b, a)_{rep}) = (b, (a, b)_{rep}) \succ (b, (b, a)_{rep})$. Let $v = ((b, (a, b)_{rep})$ and $w = (b, (b, a)_{rep})$. By (MS), there exist $s \in \mathbb{N}$ and $u, u' \in X^\infty$ such that $({}_1v_s, u) \succ ({}_1w_s, u)$ and $({}_1v_{s+1}, u') \succ ({}_1w_{s+1}, u')$. By the reflexivity of $\succeq$, $s > 1$. Note that either s or $s+1$ is odd. Without loss of generality, let s be odd. Then, we have $(b,a,b,a,b,\ldots a,b,u) \succ (b,b,a,b,a,\ldots b,a,u)$, where $a,b,a,b,\ldots a,b$ and $b,a,b,a,\ldots b,a$, respectively, are such that a,b and b,a, respectively, appear $(s-1)/2 \geq 1$ times. Since s is finite, by the repeated use of (IE) and the transitivity of $\succeq$, $(b,a,b,a,b,\ldots a,b,u) \sim (b,b,a,b,a,\ldots b,a,u)$, a contradiction.

The above three cases exhaust all possibilities. Therefore, there is no $\succeq$ satisfying (VWD), (IE) and (MS).

Remark 1. We note that the axiom of Pareto implies the axiom of very weak dominance, and that the axiom of weak dominance considered by Basu and Mitra (2007) is stronger than the axiom of very weak dominance (though they are equivalent in the presence of the axiom of intergenerational equity). It is then clear that there is no social welfare relation that simultaneously satisfies (IE), (MS) and either (P) or (WD).

Remark 2. Note that Theorem 1 holds for the case in which X contains just two different alternatives.

Remark 3. We note that a social welfare relation satisfying (MS) may or may not be representable by a social welfare function. For example, the social welfare relation $\succeq^2_{lex}$ defined below satisfies (MS), but is not representable by a social welfare function:

for all $x,y \in X^\infty$, let $x \succeq^2_{lex} y$ iff $[x_1 > y_1$ or $(x_1 = y_1$ and $x_2 \geq y_2)]$. $\succeq^2_{lex}$ satisfies (VWD), but violates (IE).

Remark 4. On the other hand, as shown by Basu and Mitra (2007), there exists a social welfare function that satisfies the axiom of partial Pareto (see Chapter 5), and (IE) when X is the set of all non-negative integers. Note that the axiom of partial Pareto is a stronger condition than (VWD). Therefore, in view of our impossibility result and the possibility result obtained in Basu and Mitra (2007), (MS) is not necessary for the representability. Together with Remark 3, it becomes clear that (MS) is an independent property from the representability of a social welfare relation by a social welfare function.

7.4 Immediate impatience and universal indifference

As shown in the last section, in the presence of (MS), any Paretian social welfare relation is not intergenerationally equitable. (MS) is a very reasonable property for a social welfare relation to satisfy. Inspired by our impossibility result, in this section, we examine the respective implications of (VWD) and (IE) under some common axioms. First, we consider two more axioms introduced below.

Independence (I) (IND(I)) For all $a \in X$, and all $x,y \in X^\infty$, $[x \succeq y \Rightarrow (a,x) \succeq (a,y)]$.

Independence (II) (IND(II)) For all $t \in \mathbb{N}$, all $x,y \in X^\infty$ and all $u_1,\ldots,u_t, v_1,\ldots,v_t \in X$, $[({}_1u_t,x) \succeq ({}_1v_t,x) \Leftrightarrow ({}_1u_t,y) \succeq ({}_1v_t,y)]$.

(IND(I)) and (IND(II)) correspond, respectively, to Diamond's axioms of (NC1) and (NC2). Similar axioms are used and discussed in Koopmans (1960) as well. As pointed out by Diamond, these two axioms reflect 'a certain type of noncomplementarity of the preferences over time or that the 'preference' over part of the time horizon are independent of the utility levels achieved in other times' (Diamond, 1965, p. 175).

Pareto principle and immediate impatience

We now explore implications of the Pareto principle in the presence of (IND(I)), (IND(II)) and (MS). The consequence is an immediate impatience result which is summarized in the following theorem.

Theorem 2. Suppose $\succeq$ satisfies (VWD), (MS), (IND(I)) and (IND(II)). Then, for all $i \in \mathbb{N}$, all $x \in X^\infty$,

$$x_i > x_{i+1} \Rightarrow x \succ x(i\,i+1).$$

Proof: Suppose $\succeq$ satisfies (VWD), (MS), (IND(I)) and (IND(II)). Let $a, b \in X$ with $a > b$. Consider $u = ((a,b)_{rep})$ and $v = ((b,a)_{rep})$. If $((b,a)_{rep}) \succeq ((a,b)_{rep})$, then, by (IND(I)), $(a,(b,a)_{rep}) \succeq (a,(a,b)_{rep})$. Note that $(a,(b,a)_{rep}) = ((a,b)_{rep})$ and $((b,a)_{rep}) = (b,(a,b)_{rep})$. The transitivity of $\succeq$ implies that $(b,(a,b)_{rep}) = ((b,a)_{rep}) \succeq (a,(a,b)_{rep})$, a contradiction to (VWD), which implies that $(a,(a,b)_{rep}) \succ ((b,a)_{rep}) = (b,(a,b)_{rep})$. Therefore, from the completeness of $\succeq$, $u = ((a,b)_{rep}) \succ ((b,a)_{rep}) = v$. By (MS), from $u = ((a,b)_{rep}) \succ ((b,a)_{rep}) = v$, there exist $z, z' \in X^\infty$ and $t \in \mathbb{N}$ such that $({}_1u_t, z) \succ ({}_1v_t, z)$ and $({}_1u_{t+1}, z) \succ ({}_1v_{t+1}, z)$. Clearly, either t or $t+1$ is an even number. Without loss of generality, let t be even. Then, we have $(a,b,\ldots,a,b,z) \succ (b,a,\ldots,b,a,z)$, where (a,b) appears $t/2$ times in $(a,b,\ldots,a,b,z)$ and (b,a) appears $t/2$ times in $(b,a,\ldots,b,a,z)$.

We now show that $(a,b,z) \succ (b,a,z)$. Suppose not, then, by the completeness of $\succeq$, we have $(b,a,z) \succeq (a,b,z)$. By using (IND(I)) twice, from $(b,a,z) \succeq (a,b,z)$, we obtain $(b,a,b,a,z) \succeq (b,a,a,b,z)$. By (IND(II)), from $(b,a,z) \succeq (a,b,z)$ and considering $(a,b,z) \in X^\infty$, we obtain $(b,a,a,b,z) \succeq (a,b,a,b,z)$. The transitivity of $\succeq$ implies that $(b,a,b,a,z) \succeq (a,b,a,b,z)$. Repeating the above procedures, if necessary, we obtain $(b,a,b,a,\ldots b,a,z) \succeq (a,b,a,b,\ldots a,b,z)$ where b,a and a,b appear $t/2$ times in their respective infinite utility streams, a contradiction with the previously established fact that $(a,b,a,b,\ldots a,b,z) \succ (b,a,b,a,\ldots b,a,z)$. Therefore, $(a,b,z) \succ (b,a,z)$. From (IND(II)), it follows that, for all $x' \in X^\infty$, $(a,b,x') \succ (b,a,x')$. Thus, we have shown that,

$$\text{for all } a, b \in X \text{ and all } x' \in X^\infty, \text{ if } a > b \text{ then } (a,b,x') \succ (b,a,x').$$

Consider any $x \in X^\infty$. If $x_i > x_{i+1}$, from the above analysis, we have $(x_i, x_{i+1}, x_{i+2}, \ldots) \succ (x_{i+1}, x_i, x_{i+2}, \ldots)$. By the repeated application of (IND(I)), we then obtain $x \succ x(i\,i+1)$.

It should be noted that Theorem 2 is not vacuous. For example, the lexicographic relation $\succeq_{lex}$ defined below:

for all $x, y \in X^{\infty}$, $x \succ_{lex} y$ if $x_1 > y_1$, or there exists $m \in N$ such that $[x_m > y_m$ and $x_j = y_j$ for all $j < m]$, and $x \sim_{lex} y$ if $x = y$,

satisfies (VWD), (MS), (IND(I)) and (IND(II)), and thus exhibits immediate impatience. Another social welfare relation, to be called the additive relation and to be denoted by $\succeq_{additive}$, defined below also satisfies (VWD), (MS), (IND(I)) and (IND(II)):

For each and every $i \in N$, there exists $f_i : X \to \mathbb{R}$ such that,
(i) for all $a \in X$ and all $j, m \in N$, $j < m \Rightarrow f_j(a) > f_m(a)$,
(ii) for all $a, b \in X$ and all $j \in N$, $a \leq b \Leftrightarrow f_j(a) \leq f_j(b)$, and
(iii) for all $x, y \in X$, $x \succeq_{additive} y \Leftrightarrow \sum_{i=1}^{\infty} f_i(x_i) \geq \sum_{i=1}^{\infty} f_i(y_i)$.

We also note that Diamond (1965) shows that if a social welfare function satisfies the (corresponding) axioms of Pareto, independence (I), independence (II), and the axiom of continuity (in the supra metric), then the social welfare function exhibits *eventual impatience* – an impatience for the first period over the tth period for all t sufficiently far in the future. Our result of Theorem 2 thus strengthens Diamond's result in two respects. First, we show that there is an immediate impatience. Second, we obtain our result without insisting on a social welfare function satisfying the axiom of continuity.

Intergenerational equity and universal indifference

In this subsection, we examine the implication of the axiom of intergenerational equity under (MS), (IND(I)) and (IND(II)). The implication is summarized in the following theorem.

Theorem 3. Suppose that $\succeq$ satisfies (IE), (MS), (IND(I)) and (IND(II)). Then, for all $x, y \in X^{\infty}$, $x \sim y$.

Proof. Suppose that $\succeq$ satisfies (IE), (MS), (IND(I)) and (IND(II)). First, we show that

$$\text{for all } a, b \in X,\ u = ((a, b)_{rep}) \sim v = ((b, a)_{rep}). \tag{7.1}$$

Let $a, b \in X$. If $((a, b)_{rep}) \succ ((b, a)_{rep})$, then (MS) implies that there exist $z, z' \in X^{\infty}$ and $t \in \mathbb{N}$ such that, $({}_1u_t, z) \succ ({}_1v_t, z)$ and $({}_1u_{t+1}, z) \succ ({}_1v_{t+1}, z)$. Either t or $t+1$ is an even number. Without loss of generality, let t be even. Then, we must have $(a, b, \ldots, a, b, z) \succ (b, a, \ldots, b, a, z)$, where (a, b) appears $t/2$ times in $(a, b, \ldots, a, b, z)$ and (b, a) appears $t/2$ times in $(b, a, \ldots, b, a, z)$.

However, on the other hand, by the repeated use of (IE) and the transitivity of $\succeq$, we have $(a, b, \ldots, a, b, z) \sim (b, a, \ldots, b, a, z)$, a contradiction. $(b, a, b, a, b, a, \ldots) \succ (a, b, a, b, a, b, \ldots)$ leads to a similar contradiction. Therefore, (1) follows from the completeness of $\succeq$.

We next show that

$$\text{for every } x \in X^\infty, (a, x) \sim (b, x). \tag{7.2}$$

To show (7.2), we first note that, by (IND(I)), from (7.1), we obtain

$$(b, (a, b)_{rep}) \sim (b, (b, a)_{rep}), \text{ and } (a, (b, a)_{rep}) \sim (a, (a, b)_{rep}). \tag{7.3}$$

From (7.1) and the first part of (7.3) and noting that $((a, b)_{rep}) = (a, (b, a)_{rep})$ and $(b, (a, b)_{rep}) = ((b, a)_{rep})$, by the transitivity of $\succeq$, we obtain

$$(a, (b, a)_{rep}) \sim (b, (b, a)_{rep}). \tag{7.4}$$

From (7.1) and the second part of (7.3) and noting that $((b, a)_{rep}) = (b, (a, b)_{rep})$ and $(a, (b, a)_{rep}) = ((a, b)_{rep})$, the transitivity of $\succeq$ implies that

$$(b, (a, b)_{rep}) \sim (a, (a, b)_{rep}). \tag{7.5}$$

If $(a, x) \succ (b, x)$, by (IND(II)), $(a, (b, a)_{rep}) \succ (b, (b, a)_{rep})$, which contradicts (7.4). If $(b, x) \succ (a, x)$, by (IND(II)), $(b, (a, b)_{rep}) \succ (a, (a, b)_{rep})$, which contradicts (7.5). Since $\succeq$ is complete, (7.2) then follows easily.

Consider any $(x_1 x_2, x')$ and $(y_1 y_2, x') \in X^\infty$. From (7.2), $(x_2, x') \sim (y_2, x')$. By (IND(I)), $(x_1, x_2, x') \sim (x_1, y_2, x')$. Similarly, $(x_1, x') \sim (y_1, x')$. By (IND(II)), $(x_1, y_2, x') \sim (y_1, y_2, x')$. Therefore, $(x_1, x_2, x') \sim (y_1, y_2, x')$ follows from the transitivity of $\succeq$. Similarly, it can be shown that

$$\text{for all } t \in \mathbb{N}, \text{ all } x' \in X^\infty, \text{all } x_1, \ldots, x_t, y_1, \ldots, y_t \in X, ({}_1x_t, x') \sim ({}_1y_t, x'). \tag{7.6}$$

Consider $x, y \in X^\infty$. If $x \succ y$, then, by (MS), there exist $t \in \mathbb{N}$ and $z \in X^\infty$ such that $(x_1, \ldots, x_t, z) \succ (y_1, \ldots, y_t, z)$, which contradicts (7.6). $y \succ x$ leads to a similar contradiction to (7.6). Therefore, for all $x, y \in X^\infty$, $x \sim y$.

It is interesting to note the following structural similarity between our Theorem 3 and Hansson's (1969) result on the constancy of a social welfare function. As shown in Hansson (1969), a social welfare function must be constant if it satisfies unrestricted domain, Hansson's independence of irrelevant alternatives, anonymity and neutrality between alternatives. (IE) is the counterpart of Hansson's anonymity. His independence of irrelevant alternatives and neutrality between alternatives are reminiscent of our (IND(I)) and (IND(II)).

7.5 Conclusion

The possibility of combining both the Pareto principle and intergenerational equity in a social welfare relation established by Svensson (1980) sounded very promising. Upon a further examination, however, the scope of *constructing* a social welfare relation that is both Paretian and intergenerationally equitable may be limited. The impossibility result of Theorem 1 shows that, under a very mild restriction on a social welfare relation, it is not possible to accommodate both Pareto and intergenerational equity in a social welfare relation. (MS) is a structural property and is reasonable. In the literature on evaluating infinite utility streams, apart from the possibility result established by Svensson (1980), there are several other possibility results obtained by various authors. For example, in the approach that uses social welfare relations to evaluate infinite utility streams Fleurbaey and Michel (2003) discuss extensions of the Ramsey principle, and Bossert, Sprumont and Suzumura (2005) and Asheim and Tungodden (2004) discuss extensions based on transfer-sensitive quasi-orderings and of leximin. In the social welfare function approach to the problem of evaluating infinite utility streams, Basu and Mitra (2003b) consider an infinite-horizon version of utilitarianism and discuss extensions of the overtaking criterion by von Weizsäcker (1965). It is fair to say that, in all those possibility results, the methods used for proving are not constructive: either the axiom of choice is invoked or Szpilrajn's (1930) lemma on extending a quasi-ordering to a complete ordering is used. These non-constructive proof methods themselves do not necessarily suggest that all the possibility results for the problem of aggregating infinite utility streams must be non-constructive. This is due to our insufficient knowledge about the necessity of Szpilrajn's lemma or the axiom of choice in proving those possibility results. On the other hand, given that all the possibility results available up to now satisfy both (VWD) and (IE), clearly, they all fail to satisfy (MS). As a future research agenda, it is then interesting to investigate the precise reason why they all fail (MS).

Given the simplicity of (MS), it can serve as a handy tool to check if a *constructed* social welfare function or social welfare relation indeed satisfies both (VWD) and (IE). Finally, our result of Theorem 3 suggests that the axiom of intergenerational equity, together with the axioms of minimal support and independence (I) and (II), puts severe restrictions on possible social welfare relations: there is just one way of ranking all infinite utility streams, which is that they all must be indifferent. On the other hand, though the axiom of weak Pareto implies immediate impatience in the presence of the axioms of minimal support, and independence (I) and (II), it offers more possibilities.

References

Asheim, G.B. and B. Tungodden (2004) 'Resolving Distributional Conflicts Between Generations', *Economic Theory*, vol. 24, pp. 221–30.

Basu, K. and T. Mitra (2003*a*) 'Aggregating Infinite Utility Streams with Intergenerational Equity: the Impossibility of being Paretian', *Econometrica*, vol. 71, pp. 1557–63.

Basu, K. and T. Mitra (2003b) 'Utilitarianism for Infinite Utility Streams: a New Welfare criterion and its Axiomatic Characterization', *Working Paper*, Department of Economics, Cornell University.

Basu, K. and T. Mitra (2007) 'Possibility Theorems for Equitably Aggregating Infinite Utility Streams', Ch. 5 in this volume.

Bossert, W., Y. Sprumont and K. Suzumura (2005) 'Ordering Infinite Utility Streams', *mimeo*, Institute of Economic Research, Hitotsubashi University.

Brock, W. (1970) 'An Axiomatic Basis for the Ramsey-Weizsäcker Overtaking Criterion', *Econometrica*, vol. 38, pp. 927–9.

Diamond, P. (1965) 'The Evaluation of Infinite Utility Streams', *Econometrica*, vol. 33, pp. 170–7.

Fleurbaey, M. and P. Michel (2003) 'Intertemporal Equity and the Extension of the Ramsey Principle', *Journal of Mathematical Economics*, vol. 39, pp. 777–802.

Hansson, B. (1969) 'Group Preferences', *Econometrica*, vol. 37, pp. 50–4.

Koopmans, T.C. (1960) 'Stationary Ordinal Utility and Impatience', *Econometrica*, vol. 28, pp. 287–309.

Svensson, L.-G. (1980) 'Equity among Generations', *Econometrica*, vol. 48, pp. 1251–6.

Szpilrajn, E. (1930) 'Sur l'extension de l'ordre partiel', *Fundamenta Mathematicae*, vol. 16, pp. 386–9.

von Weizäcker, C.C. (1965) 'Existence of Optimal Programmes of Accumulation for an Infinite Time Horizon', *Review of Economic Studies*, vol. 32, pp. 85–104.

Part III
Intergenerational Evaluations

8
Formal Welfarism and Intergenerational Equity*

Claude d'Aspremont

8.1 Introduction

Intergenerational justice is a matter that should primarily concern the present generation, since the individuals living now are those to take immediate decisions affecting generations that will be living in the future, and even in the far future, as we know, for example, from the exhaustibility of some resources or from the long-term effects of pollution such as global warming. Of course, each future generation will become 'present' at some point in time, and the reasoning followed for the present 'present generation' about intergenerational justice could be repeated at that point in time. But, to develop this reasoning, each present generation should have a representation of future generations' interests. In that respect, a simple formulation of the problem that has been extensively analyzed consists in trying to find, under equity and efficiency conditions, an ordering of the set of possible 'infinite utility streams', that is, of the set of possible infinite sequences of utility levels attached to the successive generations starting with the present generation. In such a formulation, the welfare of each generation is represented by a single utility level, as if a generation were composed of a single individual or of a cohort of identical individuals with identical allocation. Even though this formulation owes so much to Ramsey (1928), the particular ranking criterion he proposed – that of maximizing the sum of undiscounted utility – was to be rejected for its limitations (see Chakravarty, 1962), and the possibility of representing an ordering of utility streams by a collective utility function (or social welfare function), treating all generations equally, was put into serious question. Impossibility results by Koopmans (1960), Diamond (1965) have been followed by others (*e.g.* Basu and Mitra, 2003a). However,

* I thank the participants to the IEA Roundtable Meeting on Intergenerational Equity, and in particular Geir Asheim and Walter Bossert, for their fruitful comments and suggestions. Financial support from the Belgian Science Policy Office (CLIMNEG 2 and IAP programmes) is gratefully acknowledged.

if instead of looking for a social welfare function defined on infinite utility streams, one looks for a 'social welfare ordering' of these streams, then, as shown non-constructively by Svensson (1980), satisfying a strong Pareto condition and treating generations equally, in some limited sense, become possible.

Our purpose in this chapter is neither to pursue the investigation of the general possibility or impossibility issue in this formulation of the problem of justice among generations, nor to re-examine the necessity of discounting to obtain a social welfare function. Other chapters of this volume treat these questions. We want to look at the foundations of this formulation within the 'social welfare functional' approach to social choice (as introduced by Sen, 1970). In that respect it can be seen as a supplement to the overview given in d'Aspremont and Gevers[1] (2002), where the intergenerational problem is not treated.

To think about intergenerational justice in terms of infinite generational utility streams, and to look for a social welfare ordering defined on the set of such streams, presume a double reduction. The first is the classical 'welfarist' reduction, as usually defined in ethics and social choice theory, namely that 'utility' provides all the information required to construct a social evaluation rule. Of course the strength of this reduction depends on the precise interpretation given to the concept of utility (for a discussion, see Mongin and d'Aspremont, 1998). We shall not discuss various possible interpretations here and limit our analysis to the formal consequences of welfarism, an attitude that may be called 'formal welfarism'. But the welfarist reduction is not the only one to be subsumed in the infinite utility stream approach. There is also, for each generation, the aggregation of the individual utility levels at each generation into a single utility level. We shall argue that this second type of reduction is not innocuous either. Not only does it require to impose additional assumptions on the social welfare functionals taken as the primitive concept for evaluating social states. It also obliterates the relationship between the value judgments made in the social evaluation of the welfare of the set of individuals forming the present generation with that of future generations. To defend this argument, we rely on standard results in social choice theory showing the capacity for some social evaluation criteria to proliferate (Sen, 1977, and Hammond, 1979), in the sense that adopting such a criterion for a subgroup of individuals (*e.g.* the present generation) forces an ethical observer to use the same criterion for any larger group (*e.g.* any larger set of generations). The consequences of adopting some proliferating criterion in evaluating infinite utility streams, are better examined if these streams remain disaggregated at the individual level, allowing to apply the criterion to a subgroup of individuals, and, most importantly, to the present generation, and also to exploit the bulk of social choice theory as developed for the finite case. In particular, our results concerning the orderings generated respectively by the pure utilitarian rule and the Leximin rule, both having the proliferating property, are compared to the characterisations given by Basu and Mitra (2003b), Asheim and Tungodden (2004) and Bossert, Sprumont and Suzumura (2004) on different infinite-horizon extensions of these rules. To keep with the idea of an overview though, we try to be more general and derive a characterization result (the general overtaking theorem), as well as a simplified criterion, that can be associated to any rule having the proliferating property.

8.2 Welfarism for successive generations

We consider a countably infinite set of time periods, starting from the present one and denoted by $\mathbb{T} = \{0, 1, \ldots, t, \ldots\}$. Associated to each time period t there is a 'generation' made up of a finite set N_t of n_t individuals and there is a set X_t of possible social states. The set of all individuals is represented as a partition of the set of positive integers into successive generations:

$$\mathbb{N} = \{N_0, N_1, \ldots, N_t, \ldots\}.$$

We assume that $n_t \geq 2$ and $|X_t| \geq 3$, for every t. Our objective is to evaluate the respective merits of each programme of social states $x = \{x^t\} \in \mathbb{X} \equiv \chi_{t=1}^{\infty} X_t$ for an infinite future, taking into account individual evaluations. The final evaluation, of course, will have to take into account feasibility constraints, and in particular that the set of social states at some period may depend on the social states realized in previous periods. Here, however, we shall focus on the definition of general evaluation criteria applicable to any set of feasible programmes. Individuals are supposed to live a finite number of periods. For the present analysis we keep in mind two standard cases, the case where each individual lives only one period (N_t is the set of individuals living at period t), and the case where individuals live for two periods (N_t is the set of individuals born at period t) and generations overlap. In both cases, N_t is the set of individuals belonging to generation t.

To introduce intergenerational evaluation a simple and usual approach is to suppose that, for each possible programme of social states, a 'utility level' u^t can be attached to each generation t, allowing to define an infinite 'utility stream' $u = (u^0, u^1, \ldots, u^t, \ldots) \in \mathbb{R}^{\mathbb{T}}$ (with $\mathbb{R}$ the set of real numbers), and then to look for an ordering of all infinite utility streams satisfying some efficiency or equity properties. However, this approach requires us to proceed in two stages. The first stage is to construct an ordering of all infinite *individual* welfare evaluation streams (or 'utility streams'). The second is to determine by aggregation, for every infinite individual welfare evaluation stream, the welfare level attached to each generation, and then to reduce the previous ordering to an ordering defined on the set of all infinite *generational* welfare evaluation streams.

To examine these problems we start from the concept introduced by Sen (1970) of a 'Social Welfare Functional', using it both in the case of non-overlapping and in the case of overlapping generations, and then go on to the associated concept of 'Social Welfare Ordering' (the terminology fixed by Gevers, 1979).

This, formally, consists in assuming that we have an individual evaluation function (or profile for short) given by a real-valued function U defined on $\mathbb{X} \times \mathbb{N}$. That is, if $i \in N_t$ for any $t \in \mathbb{T}$, the function $U(., i)$, or U_i for short, is a real-valued function, defined on $X_t \times X_{t+1}$ in the overlapping generation case and on X_t in the non-overlapping case, and is called individual i's evaluation function. Also, for every $x \in \mathbb{X}$, the vector $U(x, .)$, or U_x for short, is a point in the infinite *individual welfare evaluation space* $\mathbb{R}^{\mathbb{N}}$ and called an infinite individual welfare evaluation stream associated to x. Given any individual evaluation function U in some admissible subset $\mathcal{D} \subset \{U | U : \mathbb{X} \times \mathbb{N} \to \mathbb{R}\}$, we are to recommend a social ranking of $\mathbb{X}$, that is an element in the set $\mathcal{R}$ of all complete and transitive binary relations on $\mathbb{X}$. A *social welfare functional* (SWFL) is a map $F : \mathcal{D} \to \mathcal{R}$ with generic image $R_U = F(U)$ (I_U and P_U denoting respectively the associated indifference and strict preference relations). If x is ranked socially at least as high as y whenever the relevant profile is U, we write $xR_U y$ (resp. $xI_U y$ or $xP_U y$ in case of indifference or strict preference).

To reduce the SWFL approach to the comparisons of generational welfare evaluation streams, we need to introduce conditions ensuring that the relative welfare (in a formal sense) of two social states can be entirely judged by comparing their respective individual evaluation vectors, independently from the other aspects of the individual profile at hand. With individuals partitioned in a sequence of successive generations, formal welfarist social evaluation may be defined and applied at different levels, according to the domain of evaluation vectors which is considered (individual or generational). Standard conditions are the following:

Domain Attainability (AD). $\forall u, v, w \in \mathbb{R}^{\mathbb{N}}$, $\exists x, y, z \in \mathbb{X}$, $\exists U \in \mathcal{D}$ *such that* $U_x = u$, $U_y = v$ *and* $U_z = w$.

This condition ensures that the set $\{r \in \mathbb{R}^{\mathbb{N}} \mid \exists x \in \mathbb{X}, \exists U \in \mathcal{D}$ such that $U_x = r\}$ fills the whole individual evaluation space $\mathbb{R}^{\mathbb{N}}$. The next condition is a Paretian principle:

Pareto Indifference (PI). $\forall U \in \mathcal{D}$, $\forall x, y \in X$, *if* $U_x = U_y$ *then* $xI_U y$.

The third condition is an Arrowian inter-profile consistency requirement imposing that the ranking of two alternatives depends only on the evaluation of these two alternatives.

Binary Independence (BI). $\forall U, V \in \mathcal{D}$, $\forall x, y \in \mathbb{X}$ *such that* $V_x = U_x$, $V_y = U_y$, $xR_U y \Leftrightarrow xR_V y$.

As it will be recalled in the next theorem, these three conditions characterize formal welfarism, that is, the possibility to define, on the individual evaluation space $\mathbb{R}^{\mathbb{N}}$, an ordering R^*, called a *Social Welfare Ordering* (SWO), which is derived from the SWFL F. Now, since we also need to introduce the possibility of aggregating the welfare of each generation, we have to add three conditions which refer explicitly to generations. The first is a separability condition on F which allows to isolate the evaluation of the welfare of each generation.

Generational Separability (GS). $\forall t \in \mathbb{T}$, $\forall U, V \in \mathcal{D}$, *if* $\forall i \in N_t$, $U_i = V_i$ *whereas,* $\forall j \in \mathbb{N} \backslash N_t$, $\forall x, y \in \mathbb{X}$, $U(x,j) = U(y,j)$ *and* $V(x,j) = V(y,j)$, *then* $R_U = R_V$.

This condition allows to derive from F, for each generation t, a SWFL F_t defined on a domain $\mathcal{D}_t$ contained in $\{U^t \mid U^t : \mathbb{X} \times N_t \to \mathbb{R}\}$ with range in $\mathcal{R}$. For each $i \in N_t$, we define $U^t(.,i)$ on the whole set $\mathbb{X}$ for simplicity of notation, but it is constant for variables outside $X_t \times X_{t+1}$ in the overlapping generation case and outside X_t in the non-overlapping case. If the relevant profile is U^t and $x \in \mathbb{X}$ is ranked socially at least as high as $y \in \mathbb{X}$, we then write $xR^t_{U^t}y$ ($xI^t_{U^t}y$ or $xP^t_{U^t}y$ in case of indifference or strict preference). Moreover, our first three conditions can be straightforwardly reformulated for each t to be applied to each derived SWFL F_t: replacing $\mathbb{R}^{\mathbb{N}}$, $\mathcal{D}$, U, V, R_U, R_V and I_U by, respectively $\mathbb{R}^{N_t}$, $\mathcal{D}_t$, U^t, V^t, $R^t_{U^t}$, $R^t_{V^t}$ and $I^t_{U^t}$ in the conditions AD, PI and BI, we get the conditions AD$_t$, PI$_t$ and BI$_t$. They will ensure the existence of an associated SWO R^*_t defined on $\mathbb{R}^{N_t}$.

Lemma 1 *Assume the SWFL F satisfies conditions AD, PI, BI and GS. Then, for every generation* $t \in \mathbb{T}$, *there is a SWFL* $F_t : \mathcal{D}_t \to \mathcal{R}$, $\mathcal{D}_t \subset \{U^t \mid U^t : \mathbb{X} \times N_t \to \mathbb{R}\}$, *which can be identified to the restriction of the SWFL F to some subset* $\bar{\mathcal{D}}_t \subset \mathcal{D}$. *The SWFL* F_t *satisfies the conditions* AD$_t$, PI$_t$ *and* BI$_t$.

Proof. Under GS, F_t can be identified to the restriction of F to the set $\bar{\mathcal{D}}_t = \{U \in \mathcal{D} : \forall j \in \mathbb{N} \backslash N_t, \forall x \in \mathbb{X}, U(x,j) = \bar{U}(x,j)\}$, for some arbitrary $\bar{U} \in \mathcal{D}$ such that $\forall x \in \mathbb{X}$, $\bar{U}_x = \bar{u} \in \mathbb{R}^{\mathbb{N}}$, so that $\mathcal{D}_t = \{U^t \mid U^t : \mathbb{X} \times N_t \to \mathbb{R}$ and $(U^t, \bar{U}^{-t}) \in \bar{\mathcal{D}}_t\}$, with $\bar{U}^{-t} = (\bar{U}_i)_{i \in \mathbb{N} \backslash N_t}$. Since F satisfies AD, PI and BI, F_t satisfies AD$_t$, PI$_t$ and BI$_t$. Indeed PI and BI should hold on $\bar{\mathcal{D}}_t$, meaning that SP$_t$ and BI$_t$ hold on $\mathcal{D}_t$. Similarly, AD$_t$ is simply the condition AD applied to all $u, v, w \in \mathbb{R}^{N_t} \times \mathbb{R}^{\mathbb{N} \backslash N_t}$ such that $u^{-t} = v^{-t} = w^{-t} = \bar{u}^{-t} \in \mathbb{R}^{\mathbb{N} \backslash N_t}$. □

Two additional conditions, both based on the derived SWFLs $\{F_t\}$, are needed in the following theorem. One is continuity. We say that a sequence $(U^{tk})_{k=1}^{\infty} \subset \mathcal{D}_t$ converges pointwise to $U^{t0} \in \mathcal{D}_t$, if $\lim_{k \to \infty} U^{tk}(x,i) = U^{t0}(x,i)$, $\forall (x,i) \in \mathbb{X} \times N_t$. The condition is then stated as

Generational Continuity (GC). $\forall t \in N$, $\forall x, y \in \mathbb{X}$, $\forall\, U^{t0} \in \mathcal{D}_t$ *and for any sequence* $(U^{tk})_{k=1}^{\infty} \subset \mathcal{D}_t$ *converging pointwise to* U^{t0}, *if* $xR^t_{U^{tk}}y$, $\forall k \geq 1$, *then* $xR^t_{U^{t0}}y$.

This property will allow to represent each SWO R^*_t by a continuous function w_t, called a *Social Welfare Function* (SWF): $\forall u^t$, $v^t \in \mathbb{R}^{N_t}$, $w_t(u^t) \geq w_t(v^t) \Longleftrightarrow u^t R^*_t v^t$.

The last condition is an 'extended Pareto' condition (Dhillon, 1998). It is a Pareto indifference condition but applied to a partition of all individuals into groups (here the generations): if all the groups are indifferent between two alternatives, then society should also be indifferent.

Generational Pareto Indifference (GPI). $\forall (U^t)_{t\in\mathbb{T}} \in \times_{t\in\mathbb{T}}\mathcal{D}_t$, $\forall x, y \in \mathbb{X}$, *if* , $xI^t_{U^t}y$, $\forall t \in \mathbb{T}$, *then* $xI_U y$ *with* $U \in \mathcal{D}$ *such that,* $\forall t \in \mathbb{T}$, $\forall i \in N_t$, $U_i = U^t_i$.

We can now prove the following result.

Theorem 1 (intergenerational welfarism) *Assume the SWFL F satisfies conditions AD, SP and BI. Then, (i) there exists a SWO* R^* *on* $\mathbb{R}^{\mathbb{N}}$ *such that, for all* $x, y \in \mathbb{X}$ *and for all* $U \in \mathcal{D}$,

$$U_x R^* U_y \Longleftrightarrow xR_U y; \tag{8.2.1}$$

also, (ii) under GS and for every generation $t \in \mathbb{T}$, *there exists a SWO* R^*_t *on* $\mathbb{R}^{N_t}$ *such that, for all* $x, y \in \mathbb{X}$ *and for all* $U^t \in \mathcal{D}_t$,

$$U^t_x R^*_t U^t_y \Longleftrightarrow xR^t_{U^t} y; \tag{8.2.2}$$

and, (iii), with GC and GPI in addition, there is, for every generation $t \in \mathbb{T}$, *a social welfare function* $w_t : \mathbb{R}^{N_t} \longrightarrow \mathbb{R}$ *such that, for all* $u^t \in \mathbb{R}^{N_t}$, $\mathbf{1}_{N_t} w_t(u^t) I^*_t u^t$ *(with* $\mathbf{1}_{N_t} = (1, \ldots, 1) \in \mathbb{R}^{N_t}$*) and there exists a SWO* $R^{\#}$ *on* $\mathbb{R}^{\mathbb{T}}$ *such that, for all* $u, v \in \mathbb{R}^{\mathbb{N}}$,

$$(w_0(u^0), w_1(u^1) \ldots, w_t(u^t), \ldots) R^{\#} (w_0(v^0), w_1(v^1) \ldots, w_t(v^t), \ldots) \Longleftrightarrow uR^* v.$$

Proof. As is well-known from the case of a finite number of individuals, under AD, the conditions PI and BI are equivalent to the condition of *Strong Neutrality* (SN): $\forall U, V \in \mathcal{D}$, $\forall x, y \in \mathbb{X}$, if there are $x', y' \in \mathbb{X}$ such that $V_{x'} = U_x$, $V_{y'} = U_y$ then $xR_U y \Leftrightarrow x'R_V y'$. The argument consists in choosing $z \in \mathbb{X} \setminus \{y, y'\}$, with z a third alternative if the two pairs coincide, and, thanks to AD, in constructing profiles U^1, U^2 and U^3 such that $U^1_x = U^1_z = u$, $U^1_y = v$, $U^2_z = u$, $U^2_y = U^2_{y'} = v$, $U^3_{x'} = U^3_z = u$ and $U^3_{y'} = v$. Applying alternately BI and PI, we get $xR_U y \Longleftrightarrow xR_{U^1} y \Longleftrightarrow zR_{U^1} y \Longleftrightarrow zR_{U^2} y \Longleftrightarrow zR_{U^2} y' \Longleftrightarrow zR_{U^3} y' \Longleftrightarrow x'R_{U^3} y' \Longleftrightarrow x'R_V y'$, and SN follows. Then, defining R^* by (8.2.1), for some $x, y \in \mathbb{X}$ and for some $U \in \mathcal{D}$, we get by SN the same relation for any $x', y' \in \mathbb{X}$ such that $V_{x'} = U_x$, $V_{y'} = U_y$, so that R^* is well-defined. Completeness and transitivity of R^* follow from AD and from the completeness and transitivity of each R_U. This proves (i). To prove (ii), we know from lemma 1 that each F_t satisfies AD_t, PI_t and BI_t. Then, repeating the same argument as in (i), we get the required SWO R^*_t defined on $\mathbb{R}^{N_t}$ (see 8.2.2). To prove (iii), we need in addition R^*_t to be continuous, *i.e.* $\forall v^t \in \mathbb{R}^{N_t}$, that the sets $\{u^t \in R^{N_t} \mid u^t R^*_t v^t\}$ and $\{u^t \in R^N \mid v^t R^*_t u^t\}$ be closed in R^{N_t}. This property is implied by GC. Indeed, if for some $v_t \in \mathbb{R}^{N_t}$ the set $\{u^t \in R^{N_t} \mid u^t R^*_t v^t\}$, say, was not closed, it would be possible to find a sequence $(u^{tk})^{\infty}_{k=1}$ in $\mathbb{R}^{N_t}$, converging to some u^{t0}, with $u^{tk} R^*_t v^t$ for all $k \geq 1$ and $v^t P^* u^{t0}$; but it would then be possible to construct a sequence $(U^{tk})^{\infty}_{k=1}$ converging pointwise to U^{t0} such that, $\forall k \geq 1$, $(U^{tk}(x, i))_{i\in N_t} = u^{tk}$ and $(U^{tk}(y, i))_{i\in N_t} = v^t$

for some $x, y \in \mathbb{X}$, implying $xR_{U^{tk}}y$, $\forall k \geq 1$, but $yP_{U^{t0}}x$, in contradiction with GC. Thus, we obtain (see *e.g.* Theorem 3 in Blackorby, Bossert and Donaldson, 2002) that, for every $t \in \mathbb{T}$, there exists a SWF $w_t : \mathbb{R}^{N_t} \to \mathbb{R}$ such that for every $u^t \in \mathbb{R}^{N_t}$, $\mathbf{1}_{N_t} w_t(u^t) I_t^* u^t$. Now, by AD, for any $u \in \mathbb{R}^{\mathbb{N}}$, there are $x, y \in \mathbb{X}$ and $U \in \mathcal{D}$ such that $U_x = u$ and $U_y = (\mathbf{1}_{N_0} w_0(u^0), \mathbf{1}_{N_1} w_1(u^1), \ldots, \mathbf{1}_{N_t} w_t(u^t), \ldots)$ so that, by GPI, if $xI^t_{U^t}y$, $\forall t \in \mathbb{T}$, then

$$uI^*(\mathbf{1}_{N_0} w_0(u^0), \mathbf{1}_{N_1} w_1(u^1), \ldots, \mathbf{1}_{N_t} w_t(u^t), \ldots).$$

So, whenever uR^*v we may write as well,

$$(\mathbf{1}_{N_t} w_t(u^t))_{t \in \mathbb{T}} R^* (\mathbf{1}_{N_t} w_t(v^t))_{t \in \mathbb{T}},$$

or taking only one representative component per generation we can write equivalently

$$(w_0(u^0), w_1(u^1), \ldots, w_t(u^t), \ldots) R^\# (w_0(v^0), w_1(v^1), \ldots, w_t(v^t), \ldots),$$

thereby defining a SWO $R^\#$ on $\mathbb{R}^{\mathbb{T}}$. □

This 'welfarism theorem', as any other welfarism theorem[2], opens the possibility to work directly in terms of SWOs and to add conditions formulated in those terms only. Of course these additional properties will, almost always, be easily translated back in SWFL terms.

8.3 Intergenerational social welfare orderings

In the previous section we have shown that, even with an infinite number of individuals partitioned into a sequence of generations, it is possible to prove a welfarism theorem, transforming the problem of finding an acceptable Social Welfare Functional into the problem of finding an appropriate Social Welfare Ordering. The theorem above even left us with two possible SWOs: R^* or $R^\#$. Formally they are completely similar. Both are orderings of all infinite evaluation streams and, if we want to get a SWO satisfying both collective efficiency and intergenerational equity conditions, the problem of constructing the ordering remains as difficult whether in the case of R^* or in the case of $R^\#$. Existence of such an ordering is difficult to establish. We shall rely on the result by Svensson (1980). Other existence results are provided in Fleurbaey and Michel (2003). In that respect, the literature studies principally the SWO $R^\#$. We want to argue that it is preferable to work with the more basic SWO R^*. A first advantage is that R^* can be derived from a SWFL under a weaker set of assumptions (using neither GS nor GC nor GPI). However, this is a formal argument and, for some results, we will have to add separability or continuity assumptions anyway. The main argument in favour of R^* is that it forces an explicit consideration of the relationship between the problem of justice among generations and the problem of justice among individuals within each generation and, primarily, within the present generation. Should not the solutions proposed for the intragenerational problem, which has up to now been the main domain of investigation in ethics and social choice, have some bearing on the solutions that should be considered for the intergenerational problem? Using standard welfarist arguments, this section answers this question positively.

To represent collective efficiency and intergenerational equity requirements, the two basic conditions that we use are the following. The first is the welfarist translation of the strong Paretian condition, Pareto indifference being trivially satisfied by construction of R^*:

Strong Pareto (SP).* $\forall u, v \in R^{\mathbb{N}}$, *if* $u \geq v$ *and* $u \neq v$, *then* uP^*v.

By $u \geq v$ we mean $u_i \geq v_i$, $\forall i \in \mathbb{N}$. This is a strong but standard efficiency condition.

As for equity, the basic condition is to keep social indifference for finite permutations of individual evaluations both within and across generations. Although this condition seems to be introducing a minimal condition of impartiality, it already excludes the use of a discount factor. A much more demanding condition would be to allow for all permutations, but then it becomes incompatible with Pareto conditions (see Lauwers and Van Liedekerke, 1995; Lauwers, 1998). Other, intermediate, impartiality conditions are studied in Fleurbaey and Michel (2003).

Finite Anonymity (FA). If* σ *is a permutation of* $M \subset \mathbb{N}$, $|M| < \infty$, *and* $u, v \in \mathbb{R}^{\mathbb{N}}$ *are such that* $u_i = v_i$, $\forall i \in \mathbb{N} \backslash M$, *and* $u_i = v_{\sigma(i)}$, $\forall i \in M$, *then* uI^*v.

The two conditions SP* and FA* are defined as properties of a SWO R^* defined in $\mathbb{R}^{\mathbb{N}}$. But they can also be defined as properties of a *quasi-ordering* R, that is, a reflexive and transitive binary relation defined in $\mathbb{R}^{\mathbb{N}}$ (resp. in a suspace $\mathbb{R}^M$, $M \subset \mathbb{N}$, or just in $\mathbb{R}^m$, $m < \infty$, with SP* and FA* then restricted to such domains), which we call a Social Welfare Quasi-ordering (SWQ) on $\mathbb{R}^{\mathbb{N}}$ (resp. on $\mathbb{R}^M$ or on $\mathbb{R}^m$). A SWO is a complete SWQ. A SWQ R is a *sub-relation* to another SWQ R' (or, equivalently, R' is an *extension* of R) if they have the same domain and for any u and v in this domain, $uPv \Longrightarrow uP'v$ and $uIv \Longrightarrow uI'v$ (P and I denoting respectively the strict preference and indifference relations associated to R). If R' is a SWO, then R' is called an *ordering extension* of R (see Bossert, Sprumont and Suzumura, 2004).

Combining the two basic conditions we obtain a well-known SWQ, first proposed by Suppes (1966) and further analysed by Sen (1970), Kolm (1972) and Hammond (1976, 1979). This version is adapted to the infinite case (see Svensson, 1980).

Definition 1 (The m-Grading Principle) The m-Grading Principle is the SWQ R^S such that: for any permutation $\sigma : M \to M$, $M \subset \mathbb{N}$ with $|M| = m$, and for any $u, v \in R^{\mathbb{N}}$, *if* $u_i \geq v_{\sigma(i)}$, $\forall i \in M$ *and* $u_i \geq v_i$, $\forall i \in \mathbb{N} \backslash M$, *then* $uR^S v$; if in addition $u_j > v_{\sigma(j)}$ for some $j \in M$, then $uP^S v$.

The m-Grading Principle is a quasi-ordering on $\mathbb{R}^{\mathbb{N}}$ that satisfies both SP* and FA*. The following result, easily adapted from lemma 3.1.1 in d'Aspremont (1985), demonstrates the capacity (at least known[3] since Sen, 1977) of the Grading Principle to proliferate through and across generations:

Lemma 2 *If a SWQ R is an extension of the* 2*-Grading Principle then it satisfies SP* and FA*. Moreover, if R satifies these two conditions, then R is an extension of the m-Grading Principle for every* $m < \infty$.

Proof. The only new argument (with respect to the finite case) is to show that the 2-Grading Principle implies SP*. But, for $u, v \in R^{\mathbb{N}}$, supposing $u \geq v$ and $u_j > v_j$ for some j, and applying the 2-Grading Principle to the pair $\{j, j+1\}$, we immediately get $uP^* v$. □

There are, of course, many quasi-orderings satisfying SP* and FA*, but all have the m-Grading Principle as sub-relation. The interest in the Grading Principle in comparing infinite utility streams comes from the following existence theorem for SWOs given by Svensson (1980, theorem 2), and based on a result due to Szpilrajn (1930) and adapted by Arrow (1951, section 3 of chapter VI).

Theorem 2 (Svensson, 1980) If a SWQ R is an extension of the m-Grading Principle for every $m < \infty$, then there exists a SWO R^* which is an ordering extension of R (and hence R^* satisfies SP* and FA*).

An important observation is that the 'proliferating' property of a SWQ, as illustrated by the Grading Principle, can be defined in general terms.

Definition 2 (proliferating sequence) For any $M \subset \mathbb{N}$, $|M| = m$, $2 \leq m < \infty$, let R_m denote a SWQ defined on $\mathbb{R}^m$ (with P_m and I_m denoting respectively the associated strict preference and indifference relations) and, for any $u \in \mathbb{R}^{\mathbb{N}}$, let $u_M \in \mathbb{R}^m$ be such that $u_M = (u_i)_{i \in M}$. Then:

(i) A SWQ R is said to *extend* R_m if, $\forall M \subset \mathbb{N}$, $|M| = m$, $\forall u, v \in \mathbb{R}^{\mathbb{N}}$, if $u_M P_m v_M$ and $u_j \geq v_j$ (resp. $u_M I_m v_M$ and $u_j = v_j$), $\forall j \in \mathbb{N} \backslash M$, then uPv (resp. uIv).

(ii) A sequence of SWQs $(R_m)_{m=2}^{\infty}$, with each R_m defined on $\mathbb{R}^m$, is said to be *proliferating* if every SWQ R defined on $\mathbb{R}^{\mathbb{N}}$ and extending R_2, also extends R_m for every $m < \infty$.

A useful property of a proliferating sequence is the following:

Lemma 3 *Consider a proliferating sequence of SWQs* $(R_m)_{m=2}^{\infty}$, *with each* R_m *defined on* $\mathbb{R}^m$. *If a SWQ R defined on* $\mathbb{R}^{\mathbb{N}}$ *extends* R_2 *and* R_2 *satisfies SP* and FA* restricted to* $\mathbb{R}^2$, *then R satisfies SP* and FA*.*

Proof. Since R_2 satisfies SP* and FA* restricted to $\mathbb{R}^2$, it is an extension of the 2-Grading Principle on $\mathbb{R}^2$ (see, e.g., lemma 3.1.1 in d'Aspremont (1985)). So R is an extension of the 2-Grading Principle on $\mathbb{R}^{\mathbb{N}}$ and the result follows from Lemma 2. □

Clearly, by Lemma 2, the sequence of SWQs corresponding to the m-Grading Principle (as defined on $\mathbb{R}^m$) is proliferating. But there are other well-known proliferating sequences. One example is based on the pure utilitarian rule.

Definition 3 (Pure *m*-Utilitarianism) The pure utilitarian SWO on $\mathbb{R}^m$, denoted R_m^{pu}, is such that: for any $u, v \in \mathbb{R}^m$, $uR_m^{pu}v$ if only if $\sum_{i=1}^m u_i \geq \sum_{i=1}^m v_i$.

We call $(R_m^{pu})_{m=2}^{\infty}$ the *pure utilitarian sequence*. We then have:

Lemma 4 *The pure utilitarian sequence* $(R_m^{pu})_{m=2}^{\infty}$ *is proliferating.*

Proof. The proof (adapted from lemma 3.3.1 in d'Aspremont, 1985) goes by induction. Suppose a SWQ R extends R_2^{pu}, $R_3^{pu}, \ldots,$ and R_m^{pu}, we want to show that it extends R_{m+1}^{pu}. Take any $M \subset \mathbb{N}$, $|M| = m \geq 2$, any $j \in \mathbb{N}\backslash M$, and any $u, v \in \mathbb{R}^{\mathbb{N}}$ such that $u_i \geq v_i$, $\forall i \in \mathbb{N}\backslash M'$, with $M' = M \cup \{j\}$. For simplicity of notation, suppose $M' = \{1, 2, \ldots, m\} \cup \{m+1\}$. Then, we can find $w \in \mathbb{R}^{\mathbb{N}}$ such that $w_i = u_i$ for $1 \neq i \neq m+1$,

$$w_1 + w_{m+1} = u_1 + u_{m+1},$$

and $w_{m+1} = v_{m+1}$. Since R extends R_2^{pu}, we have wIu. Also, since R extends R_m^{pu}, we get

$$\sum_{i=1}^{m+1} u_i = \sum_{i=1}^{m+1} w_i > \sum_{i=1}^{m+1} v_i \Longrightarrow \sum_{i=1}^{m} w_i > \sum_{i=1}^{m} v_i \Longrightarrow wPv \Longrightarrow uPv,$$

and uIv if the inequalities are replaced by equalities. □

Hence, if the pure utilitarian rule is used to evaluate the welfare of a set of individuals belonging to some generation, say the present generation, then it has to be used to evaluate the welfare of any finite set of subsequent generations.

A second example is the Leximin (the lexicographic completion of the maximin), the proliferating property of which is known since Sen (1976, 1977) and Hammond (1979). We start by defining m-Leximin using, for any $u \in \mathbb{R}^m$, the notation $u_{i(.)} \in \mathbb{R}^m$ to denote the vector with the same set of components as u but increasingly ranked.

Definition 4 (*m*-Leximin) For $2 \leq m < \infty$, the m-Leximin SWO on $\mathbb{R}^m$, denoted R_m^{lx}, is such that: for any $u, v \in \mathbb{R}^m$, $uP_m^{lx}v$ if and only if $\exists k \in \{1, 2, \ldots, m\}$ such that $u_{i(k)} > v_{i(k)}$ and $u_{i(h)} = v_{i(h)}$, for $h = 1, 2, \ldots, k-1$.

We call $(R_m^{lx})_{m=2}^{\infty}$ the *Leximin sequence*. The proliferating property of m-Leximin can be shown by a simple adaptation of the argument in d'Aspremont (1985, lemma 3.4.1).

Lemma 5 *The Leximin sequence* $(R_m^{lx})_{m=2}^{\infty}$ *is proliferating.*

Proof. The proof goes by induction. But, first, it can be verified that if a SWQ R extends R_2^{lx}, then R is an extension of the 2-Grading Principle, hence, by lemma 2, R satisfies SP* and FA*. Suppose now that R extends $R_2^{lx}, R_3^{lx}, \ldots$, and R_m^{lx}. We want to show that it extends R_{m+1}^{lx}. With FA* we can take $M = \{1, 2, \ldots, m\}$, and $m+1 \in \mathbb{N}\backslash M$, and consider any $u, v \in \mathbb{R}^{\mathbb{N}}$ such that $u_i \geq v_i$, $\forall i > m+1$, $u_1 \leq u_2 \leq \ldots \leq u_{m+1}$ and $v_1 \leq v_2 \leq \ldots \leq v_{m+1}$. Clearly uIv if and only if $u = v$. If $u_i > v_i$, for some $i > 1$, and $u_j \geq v_j$, for $j = 1, \ldots, i-1$, then uPv by applying R_k^{lx}, $2 \leq k \leq m$, or simply by SP*. It remains to be shown that uPv, whenever $u_1 > v_1$ and $u_j < v_j$, for $j = 2, \ldots, m+1$, so that $v_2 > v_1$ (otherwise we would have $u_2 > v_2$). But, then, we can find $w \in \mathbb{R}^{\mathbb{N}}$ such that $v_1 < w_1 < \min\{v_2, u_1\}$, $w_2 = u_2$ and $w_h = v_h$, $1 \neq h \neq 2$, implying uPw by application of R_m^{lx}, and wPv by application of R_2^{lx}. □

Our purpose in defining proliferating sequences is to apply Theorem 2 and to look for SWOs that are extensions of rules that are well-known and well characterized in the intragenerational case, such as Pure Utilitarianism and Leximin, in order to rank infinite utility streams. Indeed, the choice of such a rule for any generation constrains the choice of a similar rule for intergenerational comparisons. For that purpose we define a notion of 'generalized overtaking criterion', inspired by von Weizsäcker (1965) and Atsumi (1965), but applied to other rules than the pure utilitarian rule and adapted to the case where the individual evaluation of each generation is not aggregated. Such criteria consist in 'transforming the comparison of any two infinite utility paths to an infinite number of comparisons of utility paths each containing a finite number of generations' (Asheim and Tungodden, 2004). However, the present formulation of the criterion differs in two different important respects. The first is that it aims at an ordering of all infinite individual welfare evaluation streams without aggregating them into streams of each generation utility. The second is that the criterion is not only applied to any finite number of successive generations, from some point on, but to any finite set of individuals (belonging to any generation), from some size on. This leads to a more demanding criterion (implying less completeness), but which does not privilege in a specific way the present generation and the generations in the near future. The criterion treats all generations symmetrically.

Criterion 1 (general overtaking) *A SWQ R^o defined in $\mathbb{R}^{\mathbb{N}}$ is a* generalized overtaking criterion generated by a proliferating sequence of SWQs $(R_m)_{m=2}^{\infty}$ *if it extends R_2 and is such that: $\forall u, v \in \mathbb{R}^{\mathbb{N}}$, $uP^o v$ (resp. $uI^o v$) whenever $\exists \bar{m} \geq 2$ such that, $\forall M \subset \mathbb{N}$ with $|M| = m \geq \bar{m}$, $u_M P_m v_M$* (resp. $u_M I_m v_M$).

The proliferating feature of an overtaking sequence of SWOs implies that, accepting only the first element of the sequence, because we take it as an acceptable condition, we are forced to accept any subsequent element of the sequence. Hence the properties of R_2, the first accepted SWQ, are crucial. Coming back to Svensson's theorem this remark has the following important application:

Theorem 3 (general overtaking) *Suppose R^o is a generalized overtaking criterion generated by a proliferating sequence of SWQs $(R_m)_{m=2}^{\infty}$, with R_2 satisfying SP* and FA* restricted to $\mathbb{R}^2$. Then there exists a SWO R^* defined on $\mathbb{R}^{\mathbb{N}}$, an ordering extension of R^o, satisfying SP* and FA*. Moreover, a SWQ R defined on $\mathbb{R}^{\mathbb{N}}$ extends R_2 (and hence satisfies SP* and FA*) if and only if R^o is a subrelation of R.*

Proof. Consider the generalized overtaking criterion R^o generated by the sequence $(R_m)_{m=2}^{\infty}$. Because the sequence is proliferating, R^o extends R_m, $\forall m \geq 2$. Since R_2 satisfies SP* and FA* restricted to $\mathbb{R}^2$, R^o satisfies SP* and FA* and is an extension of the m-Grading Principle for every $m < \infty$ (by lemma 2 and 3). Then, by theorem 2, there exists a SWO R^* defined on $\mathbb{R}^{\mathbb{N}}$ which is an ordering extension of R^o and satisfying SP* and FA*.

Now, consider a SWQ R defined on $\mathbb{R}^{\mathbb{N}}$ extending R_2, and so satisfying SP* and FA*. Since the sequence $(R_m)_{m=2}^{\infty}$ is proliferating, R extends all subsequent R_m, $2 \leq m < \infty$. Suppose R^o, the overtaking criterion generated by this sequence, is not a subrelation of R. Then $\exists u, v \in \mathbb{R}^{\mathbb{N}}$ such that uPv (resp. uIv) does *not* hold although uP^ov (resp. uI^ov) holds, meaning that for some $\bar{m} \geq 2$, $\forall M \subset \mathbb{N}$, $|M| = m \geq \bar{m}$, $u_M P_m v_M$ (resp. $u_M I_m v_M$). Thus, $\forall M \subset \mathbb{N}$, $|M| \geq \bar{m}$, $\exists i \in \mathbb{N}\backslash M$ such that $v_i > u_i$ (resp. $v_i \neq u_i$). Otherwise there would be some $M \subset \mathbb{N}$, $|M| = m \geq \bar{m}$, such that $u_M P_m v_M$ (resp. $u_M I_m v_M$) and, $\forall i \in \mathbb{N}\backslash M$, $u_i \geq v_i$ (resp. $u_i = v_i$) implying uPv (resp. uIv), since R extends R_m. So we may select sets $M_1, \ldots, M_{\bar{m}}$ and $M' = \{i_1, \ldots, i_{\bar{m}}\}$ such that $|M_k| \geq \bar{m}$, $i_k \in \mathbb{N}\backslash M_k$, and $v_{i_k} > u_{i_k}$ for $k = 1, \ldots, \bar{m}$ (resp. either $v_{i_k} > u_{i_k}$ for all k, or $u_{i_k} > v_{i_k}$ for all k, $1 \leq k \leq \bar{m}$) and $\{i_1, \ldots, i_{k-1}\} \subset M_k$, for $k = 2, \ldots, \bar{m}$. Then by SP*, $(v_{M'}, v_{\mathbb{N}\backslash M'})P(u_{M'}, v_{\mathbb{N}\backslash M'})$ (resp. $(v_{M'}, v_{\mathbb{N}\backslash M'})P(u_{M'}, v_{\mathbb{N}\backslash M'})$ or $(u_{M'}, v_{\mathbb{N}\backslash M'})P(v_{M'}, v_{\mathbb{N}\backslash M'})$) which contradicts $u_{M'} P_{\bar{m}} v_{M'}$ (resp. $u_{M'} I_{\bar{m}} v_{M'}$), since R satisfies $R_{\bar{m}}$. □

This theorem can be applied to any generalized overtaking criterion R^o generated by a proliferating sequence of SWQs, whenever the first element in the sequence, R_2, satisfies SP* and FA* restricted to $\mathbb{R}^2$. It can in particular be applied to the pure utilitarian generalized overtaking criterion, say R^{pu}, generated by the pure utilitarian sequence $(R_m^{pu})_{m=2}^{\infty}$.

Basu and Mitra (2003b) and Asheim and Tungodden (2004) propose alternative pure utilitarian criteria. They are formulated to compare streams of generational aggregated utility streams and give precedence to the present and near futures generations. Reformulated in our framework, Basu and Mitra

(2003b) pure utilitarian SWQ, say R^{PU}, is defined by:

$$\forall u, v \in \mathbb{R}^{\mathbb{N}}, uR^{PU}v \text{ if and only if, for some } M \subset \mathbb{N},$$
$$\sum_{i \in M} u_i \geq \sum_{i \in M} v_i \text{ and } u_j \geq v_j, \forall j \in \mathbb{N}\backslash M.$$

By the above theorem, since R^{PU} satisfies SP*, FA* and extends R_2^{pu}, it is an extension of R^{pu}. Conversely, since R^{pu} satisfies SP*, FA* and is translatable (*i.e.* for any $u, v, w \in \mathbb{R}^{\mathbb{N}}$, $uR^{pu}v \Longleftrightarrow (u+w)R^{pu}(v+w)$), the argument[4] of theorem 1 in Basu and Mitra (2003b) can be used to get that R^{pu} is an extension of R_U. Asheim and Tungodden (2004) propose two alternative pure utilitarian criteria (a Catching Up and an Overtaking criterion) defined on infinite utility streams. These extensions are respectively characterized by SP*, FA*, a translation invariance condition and two alternative 'Preference Continuity' conditions. Reformulated in our framework, both criteria would satisfy R_2^{pu} so that they are extensions of R^{pu} (but less partial).

These results suggest to use the following 'simplified criterion':

Criterion 2 (simplified) *Given a sequence of SWQs* $(R_m)_{m=2}^{\infty}$*, a SWQ* R^o *is a* simplified criterion *if* $\forall u, v \in \mathbb{R}^{\mathbb{N}}$, uP^ov *(resp.* uI^ov*) if and only if, for some* $M \subset \mathbb{N}$ *with* $|M| = m$, $u_M P_m v_M$ *and* $u_j \geq v_j$ *(resp.* $u_M I_m v_M$ *and* $u_j = v_j$*),* $\forall\, j \in \mathbb{N}\backslash M$.

The following results shows that, when the sequence $(R_m)_{m=2}^{\infty}$ is proliferating, then a simplified criterion can be used equivalently to the generalized criterion generated by this sequence.

Theorem 4 *Suppose* R^o *is a generalized overtaking criterion generated by a proliferating sequence of SWQs* $(R_m)_{m=2}^{\infty}$*, with* R_2 *satisfying SP* and FA* restricted to* $\mathbb{R}^2$*. Then* R^o *is a simplified criterion.*

Proof. Let $\hat{R}$ be a simplified criterion. That is: $\forall u, v \in \mathbb{R}^{\mathbb{N}}$, $u\hat{P}v$ (resp. $u\hat{I}v$) if and only if $u_M P_m v_M$ and $u_j \geq v_j$ (resp. $u_M I_m v_M$ and $u_j = v_j$), $\forall\, j \in \mathbb{N}\backslash M$, for some $M \subset \mathbb{N}$ with $|M| = m$. We want to show that, if $\hat{R}$ extends R_2, then $\forall u, v \in \mathbb{R}^{\mathbb{N}}$, $u\hat{R}v \Longleftrightarrow uR^ov$. Since $\hat{R}$ extends R_2 and, by lemma 3, satisfies SP* and FA*, we can apply theorem 3 and so get that $\hat{R}$ is an extension of R^o. Now, suppose $u, v \in \mathbb{R}^{\mathbb{N}}$ are such that for some $M \subset \mathbb{N}$ with $|M| = m$, $u_M P_m v_M$ and $u_j \geq v_j$ (resp. $u_M I_m v_M$ and $u_j = v_j$), $\forall\, j \in \mathbb{N}\backslash M$. Since R^o extends R_m, we have uP^ov (resp. uI^ov). □

This theorem provides a much simpler characterization of the generalized overtaking criterion associated to a proliferating sequence. We have seen how it can be used in the pure utilitarian case. As another example it can be applied to the Leximin overtaking criterion, say R^{LX} generated by the Leximin sequence $(R_m^{lx})_{m=2}^{\infty}$. We have the following simplified criterion:

$$\forall u, v \in \mathbb{R}^{\mathbb{N}}, uR^{LX}v \text{ if and only if, for some } M \subset \mathbb{N}, |M| = m,$$

$$u_M R_m^{lx} v_M \text{ and } u_j \geq v_j, \forall j \in \mathbb{N}\backslash M.$$

By theorem 3, a SWQ R defined on $\mathbb{R}^{\mathbb{N}}$ is an extension of R^{LX} if and only if it satisfies R_2^{lx} (and hence satisfies SP* and FA*). As a result, R^{LX} is equivalent to the Leximin criterion defined by Bossert, Sprumont and Suzumura (2004), any extension of which is characterized by SP*, FA* and a two-person equity axiom, called Strict Equity Preference. This axiom is implied by Hammond Equity and SP* and, together with FA* implies the satisfaction of R_2^{lx} (d'Aspremont, 1985, theorem 3.4.2). To recall, Hammond Equity, as a condition on a SWQ R, is:

Hammond Equity (HE). For any pair* $\{i, j\} \subset N$ *, if* u *and* v *in* $R^{\mathbb{N}}$ *are such that* $u_j = v_j$ *for all* $j \in \mathbb{N}\backslash\{i, j\}$*, and* $v_i < u_i < u_j < v_j$*, then* uRv.

Asheim and Tungodden (2004) propose two alternative Leximin criteria the extensions of which are characterized by SP*, FA*, Hammond Equity and, again, two alternative 'Preference Continuity' conditions. In our framework, both criteria extend R_2^{lx} so that they are extensions of R^{LX} (but less partial).

There are many other examples of generalized overtaking criteria generated by proliferating sequences, since the proliferation property is a very common phenomenon among SWOs. To illustrate, we can show that pure utilitarianism can be generalized to a very large class of rules, the generalized pure utilitarian rules. The class is defined by:

Definition 5 (Generalized Pure m-Utilitarianism) The generalized pure utilitarian SWO on $\mathbb{R}^m$, denoted R_m^{gu}, is such that: for any $u, v \in \mathbb{R}^m$, $uR_m^{gu}v$ if and only if $\sum_{i=1}^m g(u_i) \geq \sum_{i=1}^m g(v_i)$, where the transformation g is a continuous and increasing real-valued function defined on $\mathbb{R}$.

There are as many rules as there are transformations g. All these rules have the proliferation property and can therefore be used to define overtaking sequences in order to generate generalized overtaking criteria, to which the overtaking theorem can be applied. For any given transformation g, we call $(R_m^{gu})_{m=2}^{\infty}$ the *generalized pure utilitarian sequence*. We then have:

Lemma 6 *The generalized pure utilitarian sequence* $(R_m^{gu})_{m=2}^{\infty}$ *is proliferating.*

Proof. The proof is similar to the one for pure utilitarianism. Suppose a SWQ R extends R_2^{gu}, $R_3^{gu}, \ldots,$ and R_m^{gu}, we want to show that it extends R_{m+1}^{gu}. Again for simplicity, take $M' = \{1, 2, \ldots, m\} \cup \{m+1\}$, and any $u, v \in \mathbb{R}^{\mathbb{N}}$ such that $u_i \geq v_i$, $\forall i \in \mathbb{N}\backslash M'$. Choose $w \in \mathbb{R}^{\mathbb{N}}$ such that $w_i = u_i$ for $1 \neq i \neq m+1$,

$$g(w_1) + g(w_{m+1}) = g(u_1) + g(u_{m+1}),$$

and $g(w_{m+1}) = g(v_{m+1})$. Since R extends R_2^{gu}, we have wIu. Also, since R extends R_m^{gu}, we get

$$\sum_{i=1}^{m+1} g(u_i) = \sum_{i=1}^{m+1} g(w_i) > \sum_{i=1}^{m+1} g(v_i) \Longrightarrow \sum_{i=1}^{m} g(w_i) > \sum_{i=1}^{m} g(v_i) \Longrightarrow wPv \Longrightarrow uPv,$$

and uIv if the inequalities are replaced by equalities. □

This class is characterized in Blackorby, Bossert and Donaldson (2002) in the finite, and in the variable, population case (for variants see Fleming, 1952, and Debreu, 1960).

The proliferating sequence $(R_m^{gu})_{m=2}^{\infty}$ can be used to characterize a generalized pure utilitarian criterion R^g for any transformation g, defined as a simplified criterion:

$$\forall u, v \in \mathbb{R}^{\mathbb{N}},\ uR^g v \text{ if and only if, for some } M \subset \mathbb{N},$$
$$\sum_{i\in M} g(u_i) \geq \sum_{i\in M} g(v_i), \text{ and } u_j \geq v_j,\ \forall j \in \mathbb{N}\backslash M.$$

There are as many criteria as there are specifications of the transformation g, and these specifications depend on the additional conditions one wishes to impose. Various axiomatized specifications have been derived in the literature (see Blackorby, Bossert and Donaldson, 2002). We shall not review them here.

8.4 Conclusion

The characterizations we have derived for various SWOs over the infinite individual welfare evaluation space, all rely on a general overtaking criterion, and its simplified version, generated by some anonymous and strongly efficient SWQ that has the proliferating property. This property leads to the definition of an associated proliferating sequence the first element of which (satisfying strong Pareto and anonymity) plays the decisive role in the characterization. This first element can sometimes be interpreted as an equity axiom. This is notably the case for the 2-leximin (the first element in the leximin proliferating sequence). But, this first element in the sequence can also be replaced by a set of other axioms that are known to be equivalent (from characterizations derived in the finite case). These characterizations, though, very often involve an 'invariance' condition, restricting the measurability and interpersonal comparability properties of individual evaluation profiles. Not all invariance conditions can be admissible. For example, already in the finite case, FA* and noncomparability, in the sense of invariance with respect to individual increasing, or simply positive affine, transformations, imply universal social indifference (excluding SP*). Such results are reviewed in d'Aspremont and Gevers (2002). The extension of an invariance condition to the infinite-horizon case is even more delicate. However, some characterizations do work as the results of Asheim and Tungodden (2004) and of Basu and Mitra (2003b), mentioned above, do show. This line of investigation should be pursued.

Here, the overtaking theorem and its applications have been presented in order to stress an important consequence of sticking to welfarism in choosing criteria for intergenerational justice: adopting a criterion in evaluating the welfare of the present generation (or a subgroup) forces use of the same criterion for any subset of subsequent generations. This can be viewed as a very restrictive consequence, since it means that the moral value judgements of the first generation have to be imposed to all the subsequent ones. But it can also be seen, more positively, as a 'time consistency' property. If the present generation solves, in some specific way, the social welfare evaluation problem, taking into consideration the social welfare evaluation of all future generations, this solution will be consistent with the solutions that future 'present generations' should advocate. This is analogous to Rawls' conception (1971, p. 287) that looks at the problem 'from the standpoint of the original position', where 'the parties do not know to which generation they belong.' Each generation solving the same problem (behind the veil of ignorance), and expecting the same solution being adopted by the other generations, has in fact correct expectations.

Notes

1. My dear friend and co-author Louis Gevers died in September 2004. His collaboration would have greatly improved the present work.
2. See *e.g.* d'Aspremont and Gevers (1977), Sen (1977), d'Aspremont (1985) and Blackorby, Bossert and Donaldson (2002).
3. The property is mentioned by Sen (1977, n 26) as suggested by Hammond as a step to derive the same property for Leximin. For a proof, see Hammond (1979).
4. Basu and Mitra (2003b) assume that utility streams belong to $[0, 1]^{\mathbb{N}}$, but the argument in their theorem 1 can be readily adapted.

References

Arrow, K. J. (1951, 2d ed.1963), *Social Choice and Individual Values* (New York: Wiley).

Asheim, G. B. and Tungodden, B. (2004) 'Resolving Distributional Conflicts between Generations', *Economic Theory*, vol. 24, pp. 221–30.

d'Aspremont, C. (1985) 'Axioms for Social Welfare Orderings,' in: Hurwicz, L., Schmeidler, D. and Sonnenschein, H. (eds), *Social Goals and Social Organizations: Essays in Memory of Elisha Pazner* (Cambridge: Cambridge University Press), pp. 19–76.

d'Aspremont, C. and Gevers, L. (1977) 'Equity and the Informational Basis of Collective Choice,' *Review of Economic Studies*, vol. 44, pp. 199–209.

d'Aspremont, C. and Gevers, L. (2002) 'Social Welfare Functionals and Interpersonal Comparability', in Arrow, K., Sen A. and Suzumura K. (eds), *Handbook of Social Choice and Welfare*, vol. I (Amsterdam: Elsevier), pp. 459–541.

Atsumi, H. (1965) 'Neoclassical Growth and the Efficient Program of Capital Accumulation', *Review of Economic Studies*, vol. 32, pp. 127–36.

Basu, K. and Mitra, T. (2003a) 'Aggregating Infinite Utility Streams with Intergenerational Equity: the Impossibility of Being Paretian', *Econometrica*, vol. 71, pp. 1557–63.

Basu, K. and Mitra, T. (2003b) 'Utilitarianism for Infinite Utility Streams: a New Welfare Criterion and its Axiomatic Characterisation', Working Paper no. 03-05, Department of Economics, Cornell University.

Blackorby, C., Bossert W. and Donaldson, D. (2002) 'Utilitarianism and the Theory of Justice', in Arrow, K., Sen A. and Suzumura K. (eds), *Handbook of Social Choice and Welfare*, vol. I (Amsterdam: Elsevier), pp. 543–96.

Bossert, W., Sprumont Y. and Suzumura K. (2004) 'The Possibility of Ordering Infinite Utility Streams', Cahier 12-2004, CIREQ.

Chakravarty, S. (1962) 'The Existence of an Optimum Savings Program', *Econometrica*, vol. 30, pp. 178–87.

Debreu, G. (1960) 'Topological Methods in Cardinal Utility Theory', in Arrow, K.J., Karlin, S., and Suppes, P. (eds), *Mathematical Methods in the Social Sciences* (Stanford: Stanford University Press), pp. 16–26.

Diamond, P. (1965) 'The Evaluation of Infinite Utility Streams', *Econometrica*, vol. 33, pp. 170–7.

Dhillon, A. (1998) 'Extended Pareto Rules and Relative Utilitarianism', *Social Choice and Welfare*, vol. 15, pp. 521–42.

Fleming, M. (1952) 'A Cardinal Concept of Welfare', *Quarterly Journal of Economics*, vol. 66, pp. 366–84.

Fleurbaey, M. and Michel, P. (2003) 'Intertemporal Equity and the Extension of the Ramsey Criterion', *Journal of Mathematical Economics*, vol. 39, pp. 777–802.

Gevers, L. (1979) 'On Interpersonal Comparability and Social Welfare Orderings', *Econometrica*, vol. 47, pp. 75–89.

Hammond, P.J. (1976) 'Equity, Arrow's Conditions, and Rawls' Difference Principle', *Econometrica*, vol. 44, pp. 793–804.

Hammond, P.J. (1979) 'Equity in Two Person Situations: Some Consequences', *Econometrica*, vol. 47, pp. 1127–35.

Kolm, S.C. (1972) *Justice et Equité* (Paris: CNRS).

Koopmans, T. C. (1960) 'Stationary Ordinal Utility and Impatience', *Econometrica*, vol. 28, pp. 287–309.

Lauwers L. (1998) 'Intertemporal Objective Functions. Strong Pareto versus Anonymity', *Mathematical Social Sciences*, vol. 35, pp. 37–55.

Lauwers, L. and Van Liedekerke, L. (1995) 'Ultraproducts and Aggregation', *Journal of Mathematical Economics*, vol. 24, pp. 217–37.

Mongin, P. and d'Aspremont, C. (1998) 'Utility Theory and Ethics', in Barberà S., Hammond P. and Seidl C. (eds), *Handbook of Utility Theory,* vol. 1: Principles (Dordrecht: Kluwer), pp. 371–481.
Ramsey, F. (1928) 'A Mathematical Theory of Savings', *Economic Journal*, vol. 38, pp. 543–59.
Rawls, J. (1971), *A Theory of Justice* (Cambridge, MA: Harvard University Press).
Sen, A. K. (1970) *Collective Choice and Social Welfare* (San Francisco: Holden-Day).
Sen, A. K. (1976) 'Welfare Inequalities and Rawlsian Axiomatics', *Theory and Decision*, vol. 7, pp. 243–62.
Sen, A. K. (1977) 'On Weights and Measures: Informational Constraints in Social Welfare Analysis', *Econometrica* vol. 45, pp. 1539–72.
Suppes, P. (1966) 'Some Formal Models of Grading Principles', *Synthèse*, vol. 6, pp. 284–306.
Svensson, L.-G. (1980) 'Equity Among Generations', *Econometrica*, vol. 48, pp. 1251–6.
Szpilrajn, E. (1930) 'Sur l'extension de l'ordre partiel', *Fundamenta Mathematicae*, vol. 16, pp. 386–9.
Weymark, J. A. (1981) 'Generalized Gini Inequality Indices', *Mathematical Social Sciences*, vol. 1, pp. 409–30.
von Weizsäcker, C. C. (1965) 'Existence of Optimal Programmes of Accumulation for an Infinite Time Horizon', *Review of Economic Studies*, vol. 32, pp. 85–104.

9
Intertemporal Social Evaluation*

Charles Blackorby, Walter Bossert and David Donaldson

9.1 Introduction

A social-evaluation functional assigns a social ranking of alternatives to each information profile in its domain. In the classical multi-profile model of social choice, profiles are restricted to welfare information: all non-welfare information is implicitly assumed to be fixed. Because of this, the conventional approach does not allow us to discern the way in which the functional makes use of non-welfare information. For that, multiple non-welfare profiles are needed. Blackorby, Bossert and Donaldson (2005a) analyze a framework in which non-welfare information may vary across information profiles. Each information profile includes a vector of individual utility functions which represent welfare information and a vector of functions which describe social and individual non-welfare information. See also Kelsey (1987) and Roberts (1980) for approaches to social choice where non-welfare information is explicitly modelled. A social-evaluation functional is welfarist if a single ordering of utility vectors, together with the utility information in a profile, is sufficient to rank all alternatives. The ordering of utility vectors is called a social-evaluation ordering. Welfarism is a consequence of three axioms: unlimited domain, Pareto indifference and binary independence of irrelevant alternatives. Unlimited domain requires that all logically possible profiles are in the domain of the functional. If everyone's well-being is the same in two alternatives for a given profile, then Pareto indifference requires the social ranking generated by that profile to declare the two alternatives equally good. This axiom is implied by a plausible property introduced by Goodin (1991). If one alternative is declared socially better than another, he suggests it should be better for at least one member of society. Binary independence is a consistency condition across profiles. If welfare and non-welfare information

* We thank Geir Asheim, Peter Vallentyne and the participants of the 2005 IEA Conference for discussions and comments. Financial support through a grant from the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

for two alternatives coincide in two profiles, it requires the ranking of the alternatives to be the same for both profiles. If anonymity, which requires individuals to be treated impartially, is added to the three welfarism axioms, anonymous welfarism results. In that case, the social-evaluation ordering is anonymous: any permutation of a utility vector is as good as the utility vector itself.

In this chapter, we employ an intertemporal structure for social evaluation. This results in a special case of the above-described model. The non-welfare information that is permitted to vary consists of individual birth dates and lengths of life in each alternative, and all other non-welfare information is assumed to be fixed. We use the term birth date merely for convenience; our approach is capable of accommodating any definition of the start of a person's life. To investigate social-evaluation orderings in an intertemporal setting, we employ a period analysis with arbitrary period length. There are multiple information profiles which provide, for each person, lifetime utility, birth date and length of life in each alternative. We assume that no person can live more than a fixed number of periods. The maximal lifetime is finite but can be arbitrarily large. We employ the standard axioms unlimited domain, binary independence of irrelevant alternatives and anonymity. Unlimited domain allows birth dates and lifetimes to be different in different alternatives. When this axiom is combined with the usual Pareto-indifference condition and binary independence, welfarism results and, as a consequence, dates of birth and lengths of life cannot affect the social ordering. To investigate the influences birth dates or lifetimes may have on social evaluation, therefore, it is necessary to consider weaker axioms. If, in any two alternatives, all individuals have the same utility levels, birth dates and lengths of life, conditional Pareto indifference requires them to be ranked as equally good. If a social-evaluation functional satisfies unlimited domain, conditional Pareto indifference, binary independence of irrelevant alternatives and anonymity, there exists an anonymous ordering of compound vectors of individual utilities, birth dates and lifetimes which, together with the information in a profile, can be used to rank the alternatives. Because non-welfare information (in particular, birth dates and lifetimes) can influence the social ordering, not all of these orderings are welfarist. The model of this paper is a fixed-population version of the variable-population framework analyzed in Blackorby, Bossert and Donaldson (1995, 1997, 1999, 2005b). However, unlike these earlier papers, we focus on the intertemporal aspect rather than the population aspect of the issue. As a consequence, we can dispense with some assumptions that are required in the variable-population case. This move necessitates some new arguments in our analysis because some of the techniques applied in our earlier work rely on the possibility of varying the population. Moreover, we provide a general result allowing for both birth dates and lifetimes to matter in addition to lifetime well-being.

In our intertemporal setting, a natural independence condition can be imposed. The axiom is called independence of the utilities of the dead and it requires the ranking of any two alternatives to be independent of the utilities of individuals whose lives are over and who had the same lifetime utilities, birth dates and lifetimes in both alternatives. Moreover, we impose an axiom we call individual intertemporal equivalence and two conditional versions of it. Individual intertemporal equivalence implies that non-trivial tradeoffs between birth dates or lengths of life on the one hand and lifetime utilities on the other are possible. The unconditional variant requires the existence of a length of life λ_0 such that, for any given compound vector of utilities, birth dates and lifetimes, and for any given birth date σ_0, it is possible to find a lifetime utility $\tilde{d}$ such that replacing any individual's utility with $\tilde{d}$, his or her birth date with σ_0 and his or her lifetime with λ_0 leads to a new vector that is as good as the original. The conditional versions of the axiom have an analogous structure but are conditional on fixed birth dates or lengths of life.

When combined with intertemporal versions of axioms such as continuity, anonymity and the above-mentioned variants of the Pareto conditions, independence of the utilities of the dead and the variant of individual intertemporal equivalence suitable for the requisite Pareto principle have remarkably strong consequences. The axioms imply independence of the utilities of all who are unconcerned, not only of those whose lives are over. In addition, depending on the version of the Pareto axiom employed, several classes of intertemporal generalized-utilitarian orderings are characterized. Intertemporal generalized utilitarianism ranks any two alternatives by comparing their total transformed utilities. Birth-date dependent generalized utilitarianism allows the transformation to depend on birth dates, and lifetime dependent generalized utilitarianism allows it to depend on lengths of life. Finally, birth-date and lifetime-dependent generalized-utilitarian orderings are such that the transformation may depend on both birth dates and lifetimes. In order to assess whether social-evaluation orderings should be sensitive to birth dates or lengths of life, consider a birth-date and lifetime dependent generalized-utilitarian ordering such that the transformation is sensitive to birth dates. In that case, there exist a lifetime-utility level, a lifetime and two different birth dates such that the transformed value of the utility level is different, given the lifetime at the two values for birth date.

Now consider an alternative in which a person has the utility level and lifetime just mentioned. We wish to compare two alternatives that differ only in the birth date of the person. Because the conditional transformation is sensitive to birth date at that utility level and lifetime, one birth date will be ranked as better than the other. The transformation is continuous in utility and, therefore, betterness is preserved for some small decrease in utility. Thus, such a ordering approves of changes in birth dates even when, in terms of well-being, no one gains and someone loses. A similar argument applies to sensitivity to lifetimes. This suggests that we should reject the conditional

Pareto axioms and, instead, opt in favour of intertemporal strong Pareto, ruling out the effects of non-welfare information. If this is done, intertemporal generalized utilitarianism results: the transformation cannot depend on birth dates or lifetimes. Another point in favour of this option is that no individual intertemporal equivalence axiom is required in the presence of intertemporal strong Pareto, which makes the resulting set of axioms used in the characterization transparent. Special forms of birth-date dependent orderings are employed frequently in economic models. In particular, orderings that are based on geometric discounting are widely used and, for that reason, we investigate them in some detail although we do not endorse them. Using a positive discount factor, the geometric birth-date dependent generalized-utilitarian orderings employ the sum of the discounted transformed utilities. An alternative class of birth-date dependent orderings is based on a linear criterion. For each utility transformation, there are two linear birth-date dependent generalized-utilitarian orderings. They are obtained by adding or subtracting the sum of the individual birth dates to or from the sum of transformed utilities.

Geometric and linear birth-date-dependent generalized utilitarianism are jointly characterized by adding a stationarity property to the axioms characterizing birth-date dependent generalized utilitarianism. Suppose that the birth date of everyone in each of two alternatives is moved forward in time by a given number of periods. Stationarity requires the ranking of the resulting alternatives to be the same. That the linear birth-date dependent orderings are included in this characterization marks another departure from the results obtained in the variable-population setting: if the population is allowed to vary, only the geometric birth-date dependent generalized-utilitarian orderings satisfy the corresponding variable-population version of the stationarity condition; see Blackorby, Bossert and Donaldson (1997, 2005b).

The concluding section of the chapter addresses two issues. The first is a discussion of our choice of domain, in particular, the possibility of assigning different birth dates to the same person in different alternatives. The second re-examines the practice of discounting and we provide arguments against the discounting of well-being and suggest that concerns for unacceptable suffering of the present generation are better addressed by imposing constraints that prevent this from happening rather than changing the social objective into an ethically undesirable one.

9.2 Welfare information and non-welfare information

We use $\mathcal{Z}_+$ to denote the set of non-negative integers and $\mathcal{Z}_{++}$ is the set of positive integers. The set of real numbers is denoted by $\mathcal{R}$ and $\mathcal{R}_{++}$ is the set of all positive reals. For $n \in \mathcal{Z}_{++}$, $\mathcal{R}^n$ is Euclidean n-space and $\mathbf{1}_n$ is the vector consisting of n ones. Our notation for vector inequalities is $\gg$, $>$, $\geq$.

Suppose there is a universal set of alternatives X with at least three elements. In order to focus on the intertemporal aspect of our investigation, we assume that the population – the set of those who ever live – is fixed and finite but note that, with a few additional assumptions, our model and the results can be reformulated in a variable-population setting; see Blackorby, Bossert and Donaldson (2005b) for a discussion. The population is denoted by $\{1,\ldots,n\}$ where $n \geq 3$. Note that the population is assumed to be finite whereas the universal set of alternatives can be (countably or uncountably) infinite or finite. At least three elements are required in X to obtain a generalization of the welfarism theorem, and the minimal number of individuals is three in order to apply a well-known separability property.

Each individual $i \in \{1,\ldots,n\}$ has a lifetime-utility function $U_i\colon X \to \mathcal{R}$ where, for all $x \in X$, $U_i(x)$ is the lifetime utility of i in alternative x. A utility (or welfare) profile is an n-tuple $U = (U_1,\ldots,U_n)$ and the set of all logically possible utility profiles is $\mathcal{U}$. For $x \in X$, we write $U(x)$ for the vector $(U_1(x),\ldots,U_n(x))$.

Time periods are indexed by non-negative integers and individuals may be born in any period after period zero. There is a finite maximal lifetime $\bar{L} \in \mathcal{Z}_{++}$ but this upper bound on the number of periods in which an individual may be alive can be arbitrarily large.

Because our objective is to examine the intertemporal aspects of social evaluation, we focus on birth dates and lengths of life as the non-welfare information that may be of relevance. For each $i \in \{1,\ldots,n\}$, $S_i\colon X \to \mathcal{Z}_+$ assigns the period just before i is born to each alternative. Analogously, $L_i\colon X \to \{1,\ldots,\bar{L}\}$ is a function that specifies i's lifetime for each alternative. Thus, in alternative $x \in X$, i is alive in periods $S_i(x) + 1,\ldots,S_i(x) + L_i(x)$. A period-before-birth-date profile is an n-tuple $S = (S_1,\ldots,S_n)$ and the set of all logically possible period-before-birth-date profiles is $\mathcal{S}$. Analogously, a length-of-life profile is an n-tuple $L = (L_1,\ldots,L_n)$ and the set of all logically possible length-of-life profiles is $\mathcal{L}$. Furthermore, we define $S(x) = (S_1(x),\ldots,S_n(x))$ and $L(x) = (L_1(x),\ldots,L_n(x))$ for all $x \in X$.

We allow birth dates to vary across alternatives, although it is often argued that birth dates are fixed. Without going into too much detail at this stage, we note that there is some variation because the duration of pregnancy is not fixed. A discussion of possible criticisms to our model and our responses are provided in the concluding section.

An information profile (a profile, for short) collects welfare information and non-welfare information in a vector $\Upsilon = (U,S,L) \in \mathcal{U} \times \mathcal{S} \times \mathcal{L}$. For $x \in X$, we write $\Upsilon(x) = (U(x),S(x),L(x))$. We define $\Omega = \mathcal{R} \times \mathcal{Z}_+ \times \{1,\ldots,\bar{L}\}$, and the set of possible compound vectors (u,s,ℓ) of utility vectors, vectors of periods before birth and vectors of lengths of life is $\Omega^n = \mathcal{R}^n \times \mathcal{Z}_+^n \times \{1,\ldots,\bar{L}\}^n$.

A social-evaluation functional assigns a social ordering of the alternatives in X to each information profile in its domain. Our model is a special case of that studied in Blackorby, Bossert and Donaldson (2005a) where non-welfare information is not restricted to birth dates and lengths of life.

Letting $\mathcal{O}$ denote the set of all orderings on X, a social-evaluation functional is a mapping $F: \mathcal{D} \to \mathcal{O}$ where $\emptyset \neq \mathcal{D} \subseteq \mathcal{U} \times \mathcal{S} \times \mathcal{L}$ is the domain of F. We write $R_\Upsilon = F(\Upsilon)$ for all $\Upsilon \in \mathcal{D}$. The asymmetric and symmetric factors of R_Υ are P_Υ and I_Υ. Many of the orderings characterized in this chapter are not welfarist – they depend on birth dates or lifetimes in addition to lifetime utilities. Nevertheless, the informational basis required for social evaluation can be simplified in the presence of some mild axioms. We introduce the axioms and state the relevant result but we do not provide a proof because it is analogous to that of the corresponding theorem in Blackorby, Bossert and Donaldson (2005a).

Our first axiom is unlimited domain. It requires the social-evaluation functional to produce a social ordering for all logically possible information profiles.

Unlimited domain: $\mathcal{D} = \mathcal{U} \times \mathcal{S} \times \mathcal{L}$.

The next axiom is a conditional version of the well-known Pareto-indifference condition from traditional social-choice theory. Our version is weaker because it requires the conclusion of the axiom only if not only welfare information but also non-welfare information is the same in two alternatives.

Conditional Pareto indifference: For all $x, y \in X$ and for all $\Upsilon \in \mathcal{D}$, if $\Upsilon(x) = \Upsilon(y)$, then $xI_\Upsilon y$.

Binary independence of irrelevant alternatives is defined as usual. If two profiles and two alternatives are such that the profiles agree on the alternatives, then the ranking of the alternatives must be the same in both profiles.

Binary independence of irrelevant alternatives: For all $x, y \in X$ and for all $\Upsilon, \bar{\Upsilon} \in \mathcal{D}$, if $\Upsilon(x) = \bar{\Upsilon}(x)$ and $\Upsilon(y) = \bar{\Upsilon}(y)$, then

$$xR_\Upsilon y \Leftrightarrow xR_{\bar{\Upsilon}} y.$$

Anonymity requires the social-evaluation functional to be independent of the labels of the individuals – everyone in society is treated equally.

Anonymity: For all $\Upsilon, \bar{\Upsilon} \in \mathcal{D}$, if there exists a bijection $\rho: \{1, \ldots, n\} \to \{1, \ldots, n\}$ such that $\Upsilon_i = \bar{\Upsilon}_{\rho(i)}$ for all $i \in \{1, \ldots, n\}$, then $R_\Upsilon = R_{\bar{\Upsilon}}$.

Anonymity is easily defended because it allows non-welfare information to matter. All that is ruled out is the claim that an individual's identity justifies special treatment, no matter what non-welfare information obtains.

An ordering R on Ω^n is anonymous if and only if, for all $(u, s, \ell) \in \Omega^n$ and for all bijections $\rho\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$,

$$((u_{\rho(1)}, \ldots, u_{\rho(n)}), (s_{\rho(1)}, \ldots, s_{\rho(n)}), (\ell_{\rho(1)}, \ldots, \ell_{\rho(n)}))I(u, s, \ell).$$

The four axioms imply that any two alternatives can be ranked by examining the lifetime utilities, the birth dates and the lifetimes obtained in the two alternatives only – no further information is required. Moreover, anonymity implies that the ranking of the compound vectors of lifetime utilities, birth dates and lengths of life is anonymous – it is insensitive with respect to the labels we give to the individuals. Because the proof is analogous to that of the version of the welfarism theorem stated in Blackorby, Bossert and Donaldson (2005a), we omit it. See d'Aspremont (2005) for an infinite-horizon variant of the welfarism theorem.

Theorem 1 *Suppose F satisfies unlimited domain. F satisfies conditional Pareto indifference, binary independence of irrelevant alternatives and anonymity if and only if there exists an anonymous social-evaluation ordering R on* Ω^n *such that, for all* $x, y \in X$ *and for all* $\Upsilon \in \mathcal{D}$,

$$xR_{\Upsilon}y \Leftrightarrow (U(x), S(x), L(x))R(U(y), S(y), L(y)).$$

The asymmetric and symmetric factors of the social-evaluation ordering R are denoted by P and I.

9.3 A preliminary result

The proof of our main result makes use of a well-known theorem in atemporal social choice. This theorem characterizes the class of generalized-utilitarian orderings by means of some plausible axioms.

Let $\overset{*}{R}$ be an ordering on $\mathcal{R}^n$ with asymmetric factor $\overset{*}{P}$ and symmetric factor $\overset{*}{I}$. The interpretation of $\overset{*}{R}$ is that of an atemporal social-evaluation ordering – it is a special case of R that depends on utility vectors only. $\overset{*}{R}$ is anonymous if and only if, for all $u \in \mathcal{R}^n$ and for all bijections $\rho\colon \{1, \ldots, n\} \to \{1, \ldots, n\}$, $(u_{\rho(1)}, \ldots, u_{\rho(n)})\overset{*}{I}u$. $\overset{*}{R}$ is a generalized-utilitarian ordering if and only if there exists a continuous and increasing function $g\colon \mathcal{R} \to \mathcal{R}$ such that, for all $u, v \in \mathcal{R}^n$,

$$u\overset{*}{R}v \Leftrightarrow \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{n} g(v_i).$$

Our first axiom is continuity. As usual, it requires that small changes in utilities do not lead to large changes in the social ranking.

Continuity: For all $u \in \mathcal{R}^n$, the sets $\{v \in \mathcal{R}^n \mid v\overset{*}{R}u\}$ and $\{v \in \mathcal{R}^n \mid u\overset{*}{R}v\}$ are closed in $\mathcal{R}^n$.

Strong Pareto requires unanimity to be respected. If everyone's well-being in a utility vector u is greater than or equal to that person's well-being in v with at least one strict inequality, u is better than v according to $\overset{*}{R}$.

Strong Pareto: For all $u, v \in \mathcal{R}^n$, if $u > v$, then $u\overset{*}{P}v$.

The final axiom of this section is a separability property. It requires $\overset{*}{R}$ to be independent of the utilities of those who are unconcerned. Suppose that a proposed social change affects only the utilities of the members of a population subgroup. Independence of the utilities of the unconcerned requires the social assessment of the change to be independent of the utility levels of people who are not in the subgroup.

Independence of the utilities of the unconcerned: For all $m \in \{1, \ldots, n-1\}$, for all $u, v \in \mathcal{R}^m$ and for all

$$\bar{u}, \bar{v} \in \mathcal{R}^{n-m}, (u, \bar{u})\overset{*}{R}(v, \bar{u}) \Leftrightarrow (u, \bar{v})\overset{*}{R}(v, \bar{v}).$$

In this definition, the individuals with utility vectors $\bar{u}$ or $\bar{v}$ are the unconcerned – they are equally well off in $(u, \bar{u})$ and $(v, \bar{u})$ and in $(u, \bar{v})$ and $(v, \bar{v})$. The axiom requires the ranking of pairs such as $(u, \bar{u})$ and $(v, \bar{u})$ to depend on the utilities of the concerned individuals only. If formulated in terms of a real-valued representation, this axiom is referred to as complete strict separability in Blackorby, Primont and Russell (1978). A corresponding separability axiom for social-evaluation functionals that depend on welfare information only can be found in d'Aspremont and Gevers (1977) where it is called separability with respect to unconcerned individuals. D'Aspremont and Gevers' separability axiom is called elimination of (the influence of) indifferent individuals in Maskin (1978) and Roberts (1980).

In the case of two individuals, the axiom is redundant because it is implied by strong Pareto but, if there are at least three individuals (an assumption we maintain throughout), it has remarkably strong consequences. When combined with continuity and strong Pareto, it characterizes generalized utilitarianism if $\overset{*}{R}$ is assumed to be anonymous. The proof of the following theorem, which employs Debreu's (1959, pp. 56–59) representation theorem and Gorman's (1968) theorem on overlapping separable sets of variables (see also Aczél (1966, p. 312) and Blackorby, Primont and Russell, (1978, p. 127)), is in Blackorby, Bossert and Donaldson (2002), for instance. Variants of the theorem can be found in Debreu (1960) and Fleming (1952).

Theorem 2 *An anonymous ordering $\overset{*}{R}$ satisfies continuity, strong Pareto and independence of the utilities of the unconcerned if and only if $\overset{*}{R}$ is a generalized-utilitarian ordering.*

9.4 Intertemporal axioms and orderings

The remaining axioms employed in this chapter are formulated for the ordering R. Given the axioms of Theorem 1, this involves no loss of generality. We introduce an intertemporal version of continuity, conditional formulations of the strong-Pareto principle and three variants of an axiom which ensures that the birth date of an individual can be changed to a specific birth date (common for all individuals) without changing the social ranking, provided the individual's lifetime utility is suitably adjusted. These conditions ensure that non-trivial trade-offs are possible.

Continuity can be formulated in an intertemporal model in a way that is analogous to its atemporal version. We require the social ranking to be continuous in lifetime utilities for any fixed pair of birth-date vectors and lifetime vectors.

Intertemporal continuity: For all $(u, s, \ell) \in \Omega^n$, the sets $\{v \in \mathcal{R}^n \mid (v, s, \ell) R(u, s, \ell)\}$ and $\{v \in \mathcal{R}^n \mid (u, s, \ell) R(v, s, \ell)\}$ are closed in $\mathcal{R}^n$.

The intertemporal version of the strong-Pareto principle has two parts. First, if each individual has the same lifetime utilities in two alternatives, they are ranked as equally good by the social ordering. Second, if everyone's utility is greater than or equal in one alternative than in another with at least one strict inequality, the former is better than the latter.

Intertemporal strong Pareto: For all $(u, s, \ell), (v, r, k) \in \Omega^n$,

(i) if $u = v$, then $(u, s, \ell) I (v, r, k)$;
(ii) if $u > v$, then $(u, s, \ell) P (v, r, k)$.

Intertemporal strong Pareto rules out the influence of non-utility information. To allow such information to matter, we introduce the following conditional versions of the axiom. The first of these applies the principle conditionally on birth dates and lifetimes, the second on birth dates only and the final one on lifetimes only.

Conditional strong Pareto: For all $(u, s, \ell), (v, r, k) \in \Omega^n$,

(i) if $s = r$, $\ell = k$ and $u = v$, then $(u, s, \ell) I (v, r, k)$;
(ii) if $s = r$, $\ell = k$ and $u > v$, then $(u, s, \ell) P (v, r, k)$.

Birth-date conditional strong Pareto: For all $(u, s, \ell), (v, r, k) \in \Omega^n$,

(i) if $s = r$ and $u = v$, then $(u, s, \ell) I (v, r, k)$;
(ii) if $s = r$ and $u > v$, then $(u, s, \ell) P (v, r, k)$.

Lifetime conditional strong Pareto: For all $(u, s, \ell), (v, r, k) \in \Omega^n$,

(i) if $\ell = k$ and $u = v$, then $(u, s, \ell) I (v, r, k)$;
(ii) if $\ell = k$ and $u > v$, then $(u, s, \ell) P (v, r, k)$.

Part (i) of conditional strong Pareto is redundant because R is reflexive. It is included in order to have the same structure as in the other strong-Pareto axioms.

The axiom individual intertemporal equivalence and its conditional counterparts ensure that, by suitably changing an individual's lifetime utility, the birth date of the person can be moved to a prespecified period without changing the social ranking. These conditions guarantee the possibility of non-degenerate trade-offs between birth dates or lifetimes and utilities. We require more notation to proceed. Let $i \in \{1, \ldots, n\}$, $(u, s, \ell) \in \Omega^n$ and $(u'_i, s'_i, \ell'_i) \in \Omega$. The vectors $v = (u_{-i}, u'_i) \in \mathcal{R}^n$, $r = (s_{-i}, s'_i) \in \mathcal{Z}^n_+$ and $k = (\ell_{-i}, \ell'_i) \in \{1, \ldots, \bar{L}\}^n$ are defined by

$$v_j = \begin{cases} u_j & \text{if } j \in \{1, \ldots, n\} \setminus \{i\}; \\ u'_j & \text{if } j = i, \end{cases}$$

$$r_j = \begin{cases} s_j & \text{if } j \in \{1, \ldots, n\} \setminus \{i\}; \\ s'_j & \text{if } j = i \end{cases}$$

and

$$k_j = \begin{cases} \ell_j & \text{if } j \in \{1, \ldots, n\} \setminus \{i\}; \\ \ell'_j & \text{if } j = i. \end{cases}$$

Individual intertemporal equivalence: There exists $\lambda_0 \in \{1, \ldots, \bar{L}\}$ such that, for all $(d, \sigma, \lambda) \in \Omega$ and for all $\sigma_0 \in \mathcal{Z}_+$, there exists $\hat{d} \in \mathcal{R}$ such that, for all $(u, s, \ell) \in \Omega^n$ and for all $i \in \{1, \ldots, n\}$,

$$((u_{-i}, \hat{d}), (s_{-i}, \sigma_0), (\ell_{-i}, \lambda_0)) I ((u_{-i}, d), (s_{-i}, \sigma), (\ell_{-i}, \lambda)).$$

Birth-date conditional individual intertemporal equivalence: There exists $\lambda_0 \in \{1, \ldots, \bar{L}\}$ such that, for all $(d, \sigma) \in \mathcal{R} \times \mathcal{Z}_+$ and for all $\sigma_0 \in \mathcal{Z}_+$, there exists $\hat{d} \in \mathcal{R}$ such that, for all $(u, s) \in \mathcal{R}^n \times \mathcal{Z}^n_+$ and for all $i \in \{1, \ldots, n\}$,

$$((u_{-i}, \hat{d}), (s_{-i}, \sigma_0), \lambda_0 \mathbf{1}_n) I ((u_{-i}, d), (s_{-i}, \sigma), \lambda_0 \mathbf{1}_n).$$

Lifetime conditional individual intertemporal equivalence: There exist $\sigma_0 \in \mathcal{Z}_+$ and $\lambda_0 \in \{1, \ldots, \bar{L}\}$ such that, for all $(d, \lambda) \in \mathcal{R} \times \{1, \ldots, \bar{L}\}$, there exists $\hat{d} \in \mathcal{R}$ such that, for all $(u, \ell) \in \mathcal{R}^n \times \{1, \ldots, \bar{L}\}^n$ and for all $i \in \{1, \ldots, n\}$,

$$((u_{-i}, \hat{d}), \sigma_0 \mathbf{1}_n, (\ell_{-i}, \lambda_0)) I ((u_{-i}, d), \sigma_0 \mathbf{1}_n, (\ell_{-i}, \lambda)).$$

The intertemporal equivalence axioms are intended to provide sufficient flexibility in order to put the requisite Pareto condition to full use. In the presence of intertemporal strong Pareto, no individual intertemporal equivalence property is required. This is the case because intertemporal strong

Pareto by itself permits us to change birth dates and lifetimes arbitrarily (without having to change individual utilities) and arrive at an alternative that is as good as the one we started out from. This, in turn, makes it possible to convert an intertemporal social-evaluation problem into an atemporal one and apply the result of section 3. If merely conditional strong Pareto is assumed rather than the full intertemporal Pareto condition, there is no guarantee that we can move people's birth dates and lifetimes and arrive at an alternative that is equally good, even if we permit utilities to change as well. This is true because conditional strong Pareto is silent whenever birth dates or lifetimes vary from one alternative to another. As a consequence, the Pareto axiom has to be supplemented with individual intertemporal equivalence so that it has sufficient power to imply the requisite properties of an atemporal social-evaluation ordering. Analogously, because birth-date conditional strong Pareto (respectively lifetime conditional strong Pareto) allows us to change lifetimes (respectively birth dates) arbitrarily and arrive at an alternative that is as good as the original, the required intertemporal equivalence property is weaker: it is sufficient to employ birth-date conditional individual intertemporal equivalence (respectively lifetime conditional individual intertemporal equivalence) because the possibility of varying lifetimes (respectively birth dates) necessary to utilize the full force of the Pareto axiom is already guaranteed by the Pareto axiom itself.

The main result of this chapter provides characterizations of generalized utilitarianism and related orderings in our intertemporal setting. In each of the definitions of the first three classes of orderings, a condition regarding the possibility of equalizing the values of the requisite transformation for different birth dates or lifetimes is imposed. This condition is required in order to ensure that the relevant individual intertemporal equivalence property is satisfied.

R is a birth-date and lifetime dependent generalized-utilitarian ordering if and only if there exist a function $h\colon \Omega \to \mathcal{R}$, continuous and increasing in its first argument, and $\lambda_0 \in \{1, \dots, \bar{L}\}$ such that $h(\mathcal{R}, \sigma_0, \lambda_0) \cap h(\mathcal{R}, \sigma, \lambda) \neq \emptyset$ for all $\sigma_0, \sigma \in \mathcal{Z}_+$ and for all $\lambda \in \{1, \dots, \bar{L}\}$ and, for all $(u, s, \ell), (v, r, k) \in \Omega^n$,

$$(u, s, \ell) R (v, r, k) \Leftrightarrow \sum_{i=1}^{n} h(u_i, s_i, \ell_i) \geq \sum_{i=1}^{n} h(v_i, r_i, k_i).$$

Analogously, R is a birth-date dependent generalized-utilitarian ordering if and only if there exists a function $f : \mathcal{R} \times \mathcal{Z}_+ \to \mathcal{R}$, continuous and increasing in its first argument, such that $f(\mathcal{R}, \sigma_0) \cap f(\mathcal{R}, \sigma) \neq \emptyset$ for all $\sigma_0, \sigma \in \mathcal{Z}_+$ and, for all $(u, s, \ell), (v, r, k) \in \Omega^n$,

$$(u, s, \ell) R (v, r, k) \Leftrightarrow \sum_{i=1}^{n} f(u_i, s_i) \geq \sum_{i=1}^{n} f(v_i, r_i). \tag{9.1}$$

R is a lifetime dependent generalized-utilitarian ordering if and only if there exist a function $e\colon \mathcal{R} \times \{1,\dots,\bar{L}\} \to \mathcal{R}$, continuous and increasing in its first argument, and $\lambda_0 \in \{1,\dots,\bar{L}\}$ such that $e(\mathcal{R},\lambda_0) \cap e(\mathcal{R},\lambda) \neq \emptyset$ for all $\lambda \in \{1,\dots,\bar{L}\}$ and, for all $(u,s,\ell),(v,r,k) \in \Omega^n$,

$$(u,s,\ell)R(v,r,k) \Leftrightarrow \sum_{i=1}^{n} e(u_i,\ell_i) \geq \sum_{i=1}^{n} e(v_i,k_i).$$

Finally, R is an intertemporal generalized-utilitarian ordering if and only if there exists a continuous and increasing function $g\colon \mathcal{R} \to \mathcal{R}$ such that, for all $(u,s,\ell),(v,r,k) \in \Omega^n$,

$$(u,s,\ell)R(v,r,k) \Leftrightarrow \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{n} g(v_i).$$

9.5 Intertemporal independence

When choices are made in a period $t \in \mathcal{Z}_{++}$, all feasible alternatives have a common history but the lifetime utilities, birth dates and lifetimes of some members of society may not be fixed. If person i is alive in period $t-1$, there may be alternatives in which i's life extends to period t and possibly beyond, whereas in other alternatives, i dies at the end of period $t-1$. This suggests that history must matter to some extent if lifetime utilities are to be taken into consideration. On the other hand, some independence property is desirable because decisions should not depend on the utilities of individuals who are long dead, for example. If, in any period, an individual's life is over in two alternatives and he or she had the same lifetime utility, birth date and lifetime in both, a plausible requirement is that the ranking of the two alternatives does not depend on the utility level of that individual. In our setting, this leads to an independence condition whose scope is limited: it applies only if the sets of those whose lives are over are identical in two alternatives and, moreover, everyone in this set had the same lifetime utility, the same birth date and the same lifetime in both.

Independence of the utilities of the dead: For all $m \in \{1,\dots,n-1\}$, for all $(u,s,\ell),(v,r,k) \in \Omega^m$, for all $(\bar{u},\bar{s},\bar{\ell}),(\bar{v},\bar{r},\bar{k}) \in \Omega^{n-m}$ and for all $t \in \mathcal{Z}_{++}$, if $\bar{s}_i + \bar{\ell}_i < t$ and $\bar{r}_i + \bar{k}_i < t$ for all $i \in \{1,\dots,n-m\}$ and $s_i + 1 \geq t$ and $r_i + 1 \geq t$ for all $i \in \{1,\dots,m\}$, then

$$((u,\bar{u}),(s,\bar{s}),(\ell,\bar{\ell}))R((v,\bar{u}),(r,\bar{s}),(k,\bar{\ell})) \Leftrightarrow ((u,\bar{v}),(s,\bar{r}),(\ell,\bar{k}))R((v,\bar{v}),(r,\bar{r}),(k,\bar{k})).$$

Independence of the utilities of the dead is a weak separability condition because it applies to individuals whose lives are over before period t only and not to all unconcerned individuals. Thus, if all generations overlap, it does not impose any restrictions. However, when combined with intertemporal strong Pareto (or with one of the conditional versions thereof and a suitable variant of individual intertemporal equivalence), the axiom becomes more powerful.

We now provide characterizations of our intertemporal variants of generalized utilitarianism by combining intertemporal continuity and independence of the utilities of the dead with the various intertemporal versions of strong Pareto and of the intertemporal-equivalence axioms.

Theorem 3

(i) *An anonymous ordering* R *satisfies intertemporal continuity, conditional strong Pareto, individual intertemporal equivalence and independence of the utilities of the dead if and only if* R *is birth-date and lifetime-dependent generalized-utilitarian.*
(ii) *An anonymous ordering* R *satisfies intertemporal continuity, birth-date conditional strong Pareto, birth-date conditional individual intertemporal equivalence and independence of the utilities of the dead if and only if* R *is birth-date dependent generalized-utilitarian.*
(iii) *An anonymous ordering* R *satisfies intertemporal continuity, lifetime conditional strong Pareto, lifetime conditional individual intertemporal equivalence and independence of the utilities of the dead if and only if* R *is lifetime dependent generalized-utilitarian.*
(iv) *An anonymous ordering* R *satisfies intertemporal continuity, intertemporal strong Pareto and independence of the utilities of the dead if and only if* R *is intertemporal generalized-utilitarian.*

Proof We provide a detailed proof of Part (i). That the birth-date and lifetime dependent generalized-utilitarian orderings satisfy intertemporal continuity, conditional strong Pareto and independence of the utilities of the dead is straightforward to verify. The existence of a $\lambda_0 \in \{1, \ldots, \bar{L}\}$ such that $h(\mathcal{R}, \sigma_0, \lambda_0) \cap h(\mathcal{R}, \sigma, \lambda)$ is non-empty for all $\sigma_0, \sigma \in \mathcal{Z}_+$ and for all $\lambda \in \{1, \ldots, \bar{L}\}$, assumed in the definition of the orderings, guarantees that individual intertemporal equivalence is satisfied.

Now suppose R is an anonymous ordering satisfying the axioms of Part (i) of the theorem statement. The proof that R is birth-date and lifetime dependent generalized-utilitarian proceeds as follows. We define an ordering $\overset{*}{R}$ on $\mathcal{R}^n$ (that is, an ordering of utility vectors) as the restriction of R that is obtained by fixing birth dates and lengths of life at specific values. We then show that $\overset{*}{R}$ satisfies the axioms of Theorem 2 and, thus, must be generalized-utilitarian. Finally, we show that all comparisons according to R can be carried out by applying $\overset{*}{R}$ to utilities that depend on birth dates and lifetimes, resulting in birth-date and lifetime dependent generalized utilitarianism.

Let λ_0 be as in the definition of individual intertemporal equivalence. Define the ordering $\overset{*}{R}$ on $\mathcal{R}^n$ by letting, for all $u, v \in \mathcal{R}^n$,

$$u\overset{*}{R}v \Leftrightarrow (u, 0\mathbf{1}_n, \lambda_0\mathbf{1}_n)R(v, 0\mathbf{1}_n, \lambda_0\mathbf{1}_n).$$

Clearly, $\overset{*}{R}$ is an anonymous ordering satisfying continuity and strong Pareto. The last remaining property of $\overset{*}{R}$ to be established is independence of the utilities of the unconcerned. Let $m \in \{1, \ldots, n-1\}$, $u, v \in \mathcal{R}^m$ and $\bar{u}, \bar{v} \in \mathcal{R}^{n-m}$. By repeated application of individual intertemporal equivalence, there exist $\hat{u}, \hat{v} \in \mathcal{R}^m$ such that

$$((\hat{u}, \bar{u}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n)\ I((u, \bar{u}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n), \tag{9.2}$$

$$((\hat{v}, \bar{u}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n)\ I((v, \bar{u}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n), \tag{9.3}$$

$$((\hat{u}, \bar{v}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n)\ I((u, \bar{v}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n) \tag{9.4}$$

and

$$((\hat{v}, \bar{v}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n)I((v, \bar{v}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n). \tag{9.5}$$

(9.2) and (9.3) together imply

$$\begin{aligned}&((u, \bar{u}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n)\ R\ ((v, \bar{u}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n) \Leftrightarrow\\ &((\hat{u}, \bar{u}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n)\ R\ ((\hat{v}, \bar{u}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n).\end{aligned} \tag{9.6}$$

By independence of the utilities of the dead,

$$\begin{aligned}&((\hat{u}, \bar{u}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n)\ R\ ((\hat{v}, \bar{u}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n) \Leftrightarrow\\ &((\hat{u}, \bar{v}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n)\ R\ ((\hat{v}, \bar{v}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n).\end{aligned} \tag{9.7}$$

(9.4) and (9.5) together imply

$$\begin{aligned}&((\hat{u}, \bar{v}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n)\ R\ ((\hat{v}, \bar{v}), (\bar{L}\mathbf{1}_m, 0\mathbf{1}_{n-m}), \lambda_0\mathbf{1}_n) \Leftrightarrow\\ &((u, \bar{v}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n)\ R((v, \bar{v}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n).\end{aligned} \tag{9.8}$$

Combining (9.6), (9.7) and (9.8), we obtain

$$((u, \bar{u}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n)\ R((v, \bar{u}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n) \Leftrightarrow ((u, \bar{v}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n)\ R((v, \bar{v}), 0\mathbf{1}_n, \lambda_0\mathbf{1}_n).$$

By definition of $\overset{*}{R}$, this is equivalent to

$$(u,\bar{u})\overset{*}{R}(v,\bar{u}) \Leftrightarrow (u,\bar{v})\overset{*}{R}(v,\bar{v})$$

which establishes that independence of the utilities of the unconcerned is satisfied.

By Theorem 2, $\overset{*}{R}$ is generalized-utilitarian and, thus, there exists a continuous and increasing function $g\colon \mathcal{R} \to \mathcal{R}$ such that

$$u\overset{*}{R}v \Leftrightarrow \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{m} g(v_i)$$

for all $u, v \in \mathcal{R}^n$. Define the function $\bar{h}\colon \Omega \to \mathcal{R}$ by

$$\bar{h}(d,\sigma,\lambda) = \gamma \Leftrightarrow (d,\sigma,\lambda)I(\gamma,0,\lambda_0)$$

for all $(d,\sigma,\lambda) \in \Omega$ and for all $\gamma \in \mathcal{R}$. This function is well-defined because R satisfies individual intertemporal equivalence.

Consider any $(u,s,\ell),(v,r,k) \in \Omega^n$. By repeated application of individual intertemporal equivalence,

$$\left((\bar{h}(u_i,s_i,\ell_i))_{i=1}^{n}, 0\mathbf{1}_n, \lambda_0\mathbf{1}_n\right)I(u,s,l)$$

and

$$\left((\bar{h}(v_i,r_i,k_i))_{i=1}^{n}, 0\mathbf{1}_n, \lambda_0\mathbf{1}_n\right)I(v,r,k).$$

Therefore,

$$(u,s,\ell)R(v,r,k) \Leftrightarrow \left((\bar{h}(u_i,s_i,\ell_i))_{i=1}^{n}, 0\mathbf{1}_n, \lambda_0\mathbf{1}_n\right) R \left((\bar{h}(v_i,r_i,k_i))_{i=1}^{n}, 0\mathbf{1}_n, \lambda_0\mathbf{1}_n\right)$$

$$\Leftrightarrow (\bar{h}(u_i,s_i,\ell_i))_{i=1}^{n}\overset{*}{R}(\bar{h}(v_i,r_i,k_i))_{i=1}^{n} \Leftrightarrow \sum_{i=1}^{n} g(\bar{h}(u_i,s_i,\ell_i)) \geq \sum_{i=1}^{n} g(\bar{h}(v_i,r_i,k_i)).$$

Letting $h = g \circ \bar{h}$ (where $\circ$ denotes function composition), it follows that

$$(u,s,\ell)R(v,r,k) \Leftrightarrow \sum_{i=1}^{n} h(u_i,s_i,\ell_i) \geq \sum_{i=1}^{n} h(v_i,r_i,k_i).$$

That h satisfies $h(\mathcal{R},\sigma_0,\lambda_0) \cap h(\mathcal{R},\sigma,\lambda) \neq \emptyset$ for all $\sigma_0, \sigma \in \mathcal{Z}_+$ and for all $\lambda \in \{1,\ldots,\bar{L}\}$ follows from the definitions of $\bar{h}$ and h. This completes the proof of Part (i).

The proofs of Parts (ii) through (iv) of the theorem are analogous. Because R is independent of lifetimes in Part (ii) and independent of birth dates in Part (iii), the corresponding weakenings of individual intertemporal equivalence are sufficient for the characterization results. In Part (iv), the equivalence axioms can be dispensed with altogether because, by intertemporal strong Pareto, the ordering cannot depend on either birth dates or lifetimes.

As mentioned after the definitions of the axioms, individual intertemporal equivalence and its conditional counterparts are intended to complement the Pareto conditions of the first three parts of the theorem. The role of these axioms is to endow the requisite Pareto principle with sufficient power to conclude that independence of the utilities of the dead implies that a suitably defined atemporal social-evaluation ordering satisfies independence of the utilities of the unconcerned. Conversely, the equivalence axioms are employed when going back from the atemporal ordering to the resulting intertemporal ordering. It should be noted that no intertemporal equivalence axiom is required in the last part of the theorem – intertemporal strong Pareto by itself is sufficient to conclude that the atemporal ordering in the proof satisfies independence of the utilities of unconcerned.

9.6 Geometric and linear discounting

In many intertemporal models, geometric discounting is employed. According to geometric discounting, the transformed lifetime utility of each person $i \in \{1, \ldots, n\}$ is multiplied by δ^{s_i}, where $\delta \in \mathcal{R}_{++}$ is a fixed discount factor. Clearly, higher values of δ are associated with higher relative weights given to future generations. A value of $\delta > 1$ corresponds to 'upcounting' – putting higher weights on later generations than on earlier generations. If $\delta = 1$, there is neither discounting nor upcounting. The most common case occurs when $\delta < 1$ – the later someone is born, the lower the weight attached to this person's lifetime utility.

The ordering R is geometric birth-date dependent generalized-utilitarian if and only if there exist a continuous and increasing function $g\colon \mathcal{R} \to \mathcal{R}$ and a constant $\delta \in \mathcal{R}_{++}$ such that, for all $(u, s, \ell), (v, k, r) \in \Omega^n$,

$$(u, s, \ell)\ R\ (v, r, k) \Leftrightarrow \sum_{i=1}^{n} \delta^{s_i} g(u_i) \geq \sum_{i=1}^{n} \delta^{r_i} g(v_i).$$

An alternative class of birth-date dependent orderings uses information on birth dates in a linear fashion. R is linear birth-date dependent generalized-utilitarian if and only if there exist a continuous and increasing function $g\colon \mathcal{R} \to \mathcal{R}$ and a constant $\beta \in \{-1, 1\}$ such that, for all $(u, s, \ell), (v, k, r) \in \Omega^n$,

$$(u, s, \ell)\ R\ (v, r, k) \Leftrightarrow \sum_{i=1}^{n} g(u_i) + \beta \sum_{i=1}^{n} s_i \geq \sum_{i=1}^{n} g(v_i) + \beta \sum_{i=1}^{n} r_i.$$

In the case of $\beta = 1$, *ceteris paribus*, later births are considered better. If $\beta = -1$, the earlier people are born, the better the corresponding alternative is (provided that lifetime utilities are the same).

We do not endorse these two classes (or any other birth-date dependent orderings) because we believe that intertemporal strong Pareto is a compelling axiom and, thus, only lifetime-utility information should matter. However, because of their link to a well-known stationarity property (and because of the important status geometric discounting enjoys in intertemporal economic models), we provide a characterization of these orderings to illustrate their properties and the ethical judgments underlying their use. Stationarity requires that the ranking of any two elements of Ω^n is unchanged if, *ceteris paribus*, the birth date of everyone is moved into the future by any number of periods in both. This is a variant of one of the most commonly used restrictions on multi-period social-evaluation orderings, and it implies that we do not attach significance to the way time periods are numbered.

Stationarity: For all $(u, s, \ell), (v, k, r) \in \Omega^n$ and for all $\tau \in \mathcal{Z}_+$,

$$(u, s + \tau \mathbf{1}_n, \ell) \text{ R } (v, r + \tau \mathbf{1}_n, k) \Leftrightarrow (u, s, \ell) R (v, r, k).$$

This stationarity axiom differs from that employed by Koopmans (1960) in the context of infinite-horizon social evaluation. Because birth dates are not allowed to vary in Koopmans' model, his property can be interpreted as requiring that social evaluation does not depend on what is considered the first period, that is, it is independent of the definition of a particular calendar. In our model, it is crucial that birth dates are allowed to vary and, as a consequence, the formulation of the stationarity axiom we use seems appropriate for our purposes.

We now obtain a joint characterization of geometric and linear birth-date dependent generalized utilitarianism by adding stationarity to the axioms characterizing birth-date dependent generalized utilitarianism.

Theorem 4 *An anonymous ordering* R *satisfies intertemporal continuity, birth-date conditional strong Pareto, birth-date conditional individual intertemporal equivalence, independence of the existence of the dead and stationarity if and only if* R *is geometric or linear birth-date dependent generalized-utilitarian.*

Proof: That geometric and linear birth-date dependent generalized utilitarianism satisfy the axioms of the theorem statement is straightforward to verify. Conversely, suppose R satisfies the required axioms. By Part (ii) of Theorem 3, R is birth-date dependent generalized-utilitarian. Stationarity implies that, for all $(u, \ell), (v, k) \in \mathcal{R}^n \times \{1, \ldots, \bar{L}\}^n$ and for all $\sigma, \tau \in \mathcal{Z}_+$,

$$(u, (\sigma + \tau)\mathbf{1}_n, \ell) \; R \; (v, (\sigma + \tau)\mathbf{1}_n, k) \Leftrightarrow (u, \sigma \mathbf{1}_n, \ell) R (v, \sigma \mathbf{1}_n, k).$$

Letting f be as in the definition of birth-date dependent generalized utilitarianism, this is equivalent to

$$\sum_{i=1}^{n} f(u_i, \sigma + \tau) \geq \sum_{i=1}^{n} f(v_i, \sigma + \tau) \Leftrightarrow \sum_{i=1}^{n} f(u_i, \sigma) \geq \sum_{i=1}^{n} f(v_i, \sigma).$$

Thus, for each $\tau \in \mathcal{Z}_+$, there exists an increasing function $\varphi_\tau\colon \mathcal{R} \to \mathcal{R}$ such that

$$\sum_{i=1}^{n} f(u_i, \sigma + \tau) = \varphi_\tau \left(\sum_{i=1}^{n} f(u_i, \sigma) \right)$$

for all $u \in \mathcal{R}^n$ and for all $\sigma, \tau \in \mathcal{Z}_+$. For each $\sigma \in \mathcal{Z}_+$, define the function $\bar{g}_\sigma \colon f(\mathcal{R}, \sigma) \to \mathcal{R}$ by

$$\bar{g}_\sigma(\gamma) = z \Leftrightarrow f(z, \sigma) = \gamma$$

for all $\gamma \in f(\mathcal{R}, \sigma)$ and for all $z \in \mathcal{R}$, that is, $\bar{g}_\sigma$ is the inverse of f with respect to its first argument for the fixed value σ of its second argument. Now let $x_i = f(u_i, \sigma)$ for all $i \in \{1, \ldots, n\}$ and $\bar{f}(\gamma, \sigma + \tau) = f(\bar{g}_\sigma(\gamma), \sigma + \tau)$ to obtain the functional equation

$$\sum_{i=1}^{n} \bar{f}(x_i, \sigma + \tau) = \varphi_\tau \left(\sum_{i=1}^{n} x_i \right).$$

This is a Pexider equation in the variables $x_1, \ldots, x_n$ the solution of which satisfies (see Aczél (1966, p. 142))

$$\bar{f}(\gamma, \sigma + \tau) = a(\tau)\gamma + b(\tau)$$

with functions $a\colon \mathcal{Z}_+ \to \mathcal{R}$ and $b\colon \mathcal{Z}_+ \to \mathcal{R}$ which do not depend on σ because φ_τ does not. Substituting back into the definition of $\bar{f}$, we obtain

$$f(z, \sigma + \tau) = a(\tau) f(z, \sigma) + b(\tau) \tag{9.9}$$

for all $z \in \mathcal{R}$ and for all $\sigma, \tau \in \mathcal{Z}_+$. Setting $\sigma = 0$, it follows that

$$f(z, \tau) = a(\tau) f(z, 0) + b(\tau) \tag{9.10}$$

for all $z \in \mathcal{R}$ and for all $\tau \in \mathcal{Z}_+$. Therefore,

$$f(z, \sigma + \tau) = a(\sigma + \tau) f(z, 0) + b(\sigma + \tau) \tag{9.11}$$

for all $z \in \mathcal{R}$ and for all $\sigma, \tau \in \mathcal{Z}_+$. Substituting (9.10) into (9.9), we obtain

$$f(z, \sigma + \tau) = a(\tau)[a(\sigma)f(z, 0) + b(\sigma)] + b(\tau) \tag{9.12}$$

for all $z \in \mathcal{R}$ and for all $\sigma, \tau \in \mathcal{Z}_+$. Combining (9.11) and (9.12), it follows that

$$a(\sigma + \tau)f(z, 0) + b(\sigma + \tau) = a(\tau)[a(\sigma)f(z, 0) + b(\sigma)] + b(\tau)$$

or, equivalently,

$$[a(\sigma + \tau) - a(\tau)a(\sigma)]f(z, 0) = a(\tau)b(\sigma) + b(\tau) - b(\sigma + \tau).$$

Because f is increasing in its first argument and the right side of this equation is independent of z, it must be the case that both sides are identically zero which requires

$$a(\sigma + \tau) = a(\tau)a(\sigma) \tag{9.13}$$

and

$$a(\tau)b(\sigma) + b(\tau) = b(\sigma + \tau) \tag{9.14}$$

for all $\sigma, \tau \in \mathcal{Z}_+$. Setting $\sigma = 0$ and $\tau = 1$ in (9.13), it follows that $a(0) = 1$. Thus, defining $\delta = a(1)$, repeated application of (9.13) implies

$$a(\sigma) = \delta^\sigma \tag{9.15}$$

for all $\sigma \in \mathcal{Z}_+$.

Suppose first that $\delta \neq 1$. Using (9.15), (9.14) implies

$$b(\sigma + \tau) = \delta^\tau b(\sigma) + b(\tau)$$

for all $\sigma, \tau \in \mathcal{Z}_+$. Interchanging the roles of σ and τ in this equation, it follows that

$$b(\sigma + \tau) = \delta^\sigma b(\tau) + b(\sigma)$$

and, therefore, we must have

$$\delta^\tau b(\sigma) + b(\tau) = \delta^\sigma b(\tau) + b(\sigma)$$

for all $\sigma, \tau \in \mathcal{Z}_+$. Setting $\tau = 1$, this implies

$$b(\sigma) = \frac{b(1)}{1 - \delta}(1 - \delta^\sigma) \tag{9.16}$$

for all $\sigma \in \mathcal{Z}_+$. By (9.10),

$$f(z,\sigma) = a(\sigma)f(z,0)+b(\sigma) = \delta^\sigma f(z,0)+\frac{b(1)}{1-\delta}(1-\delta^\sigma) = \delta^\sigma(f(z,0) = \frac{b(1)}{1-\delta})+\frac{b(1)}{1-\delta}$$

for all $z \in \mathcal{R}$ and for all $\sigma \in \mathcal{Z}_+$. Defining

$$g(z) = f(z,0) = \frac{b(1)}{1-\delta}$$

for all $z \in \mathcal{R}$, it follows that

$$f(z,\sigma) = \delta^\sigma g(z) + \frac{b(1)}{1-\delta}$$

for all $z \in \mathcal{R}$ and for all $\sigma \in \mathcal{Z}_+$. Substituting into (9.1), we obtain

$$\begin{aligned}(u,s,\ell)R(v,r,k) \Leftrightarrow & \sum_{i=1}^{n}\left[\delta^{s_i}g(u_i) + \frac{b(1)}{1-\delta}\right] \geq \sum_{i=1}^{n}\left[\delta^{r_i}g(v_i) + \frac{b(1)}{1-\delta}\right] \\ & \Leftrightarrow \sum_{i=1}^{n}\delta^{s_i}g(u_i) \geq \sum_{i=1}^{n}\delta^{r_i}g(v_i)\end{aligned}$$

for all $(u,s,\ell),(v,k,r) \in \Omega^n$. By birth-date conditional strong Pareto, δ must be positive and R is geometric birth-date dependent generalized-utilitarian.

Now suppose $\delta = 1$. Thus, $a(\sigma) = 1$ for all $\sigma \in \mathcal{Z}_+$ and (9.14) implies

$$b(\sigma + \tau) = b(\sigma) + b(\tau)$$

for all $\sigma, \tau \in \mathcal{Z}_+$. Using a straightforward induction argument, we obtain

$$b(\sigma) = b(1)\sigma$$

for all $\sigma \in \mathcal{Z}_+$. Note that, using l'Hôpital's rule, this is the limiting case of (9.16) as δ approaches one. By (9.10),

$$f(z,\sigma) = a(\sigma)f(z,0) + b(\sigma) = f(z,0) + b(1)\sigma$$

for all $z \in \mathcal{R}$ and for all $\sigma \in \mathcal{Z}_+$.

Consider first the case $b(1) = 0$. Defining

$$g](z) = f(z,0)$$

for all $z \in \mathcal{R}$ and substituting into (9.1), we obtain

$$(u, s, \ell)R(v, r, k) \Leftrightarrow \sum_{i=1}^{n} g(u_i) \geq \sum_{i=1}^{n} g(v_i)$$

for all $(u, s, \ell), (v, k, r) \in \Omega^n$. Thus, R is intertemporal generalized utilitarianism, a special case of geometric birth-date dependent generalized utilitarianism with a parameter value of $\delta = 1$.

Now consider the case $b(1) \neq 0$. Defining

$$g(z) = \frac{f(z, 0)}{|b(1)|}$$

for all $z \in \mathcal{R}$, it follows that

$$f(z, \sigma) = |b(1)|g(z) + |b(1)|\text{sign}(b(1))\sigma$$

for all $z \in \mathcal{R}$ and for all $\sigma \in \mathcal{Z}_+$. Letting $\beta = \text{sign}(b(1)) \in \{-1, 1\}$ and substituting into (9.1), we obtain

$$(u, s, \ell)R(v, r, k) \Leftrightarrow |b(1)| \left[\sum_{i=1}^{n} g(u_i) + \beta \sum_{i=1}^{n} s_i\right] \geq |b(1)| \left[\sum_{i=1}^{n} g(v_i) + \beta \sum_{i=1}^{n} r_i\right]$$
$$\Leftrightarrow \sum_{i=1}^{n} g(u_i) + \beta \sum_{i=1}^{n} s_i \geq \sum_{i=1}^{n} g(v_i) + \beta \sum_{i=1}^{n} r_i$$

for all $(u, s, \ell), (v, k, r) \in \Omega^n$ and, thus, R is linear birth-date dependent generalized utilitarian.

9.7 Concluding remarks

A possible objection to the way we model non-welfare information (in particular, information on birth dates) is the claim that an individual's birth date is fixed and, thus, that our domain which allows us to assign any birth date to an individual is too large. While it is true that a person cannot be born at a completely arbitrary time, his or her birth date may vary over several months because the duration of pregnancy is not fixed. Given the axioms employed in this chapter, this possibility is sufficient for our results. There is another possible criticism, namely, that any change in birth date – even if it is only a matter of a single period – does not allow us to treat the individual born in a period as the same individual as a person born in a later period instead. This position articulates the view that a person's birth date is a characteristic of that person and cannot be changed without changing the person. In a variable-population setting, there is an alternative to the approach that we have chosen that can accommodate this criticism. If each individual is assumed to have a fixed birth date, our axioms intertemporal strong Pareto and anonymity can be replaced with a single axiom that extends the Pareto condition so that it applies anonymously to alternatives with the same population size.

If two alternatives have the same population size and the list of utilities in the first is a permutation of the list of utilities in the second, anonymous intertemporal strong Pareto implies that the two alternatives are ranked as equally good. This move is analogous to Suppes's (1966) grading principle. Similar combined axioms correspond to the other Pareto axioms. We think that the combined axioms have strong ethical appeal and, as a consequence, can serve as a convincing defence against the objection. An argument that is sometimes made in favour of discounting is that very large sacrifices by those presently alive may be justified by larger gains to people who will exist in the distant future only. If these sacrifices are considered too demanding, discounting might be proposed to alleviate the negative effects on the generations that live earlier.

However, this argument rests on the false claim that discounting necessarily increases the well-being of the present generation. To see that the claim is not true, consider a three-person society and suppose two alternatives x and y are such that person i is born in period i for all $i \in \{1, 2, 3\}$. In x, utility levels are $u_1 = 28$, $u_2 = 4$ and $u_3 = 44$ and, in y, lifetime utilities are $u_1 = u_2 = u_3 = 24$. If intertemporal generalized utilitarianism with the identity mapping as the transformation is used to evaluate the alternatives, x is better than y and the utility level of person 1, who represents the present generation, is 28. Alternatively suppose that geometric birth-date dependent generalized utilitarianism with the identity mapping and a discount factor of $\delta = 1/2$ is used instead. In that case, the sums of discounted utilities are $28 + 2 + 11 = 41$ for x and $24 + 12 + 6 = 42$ for y, so y is better and person 1's utility is 24. As a result of discounting, the present generation is worse off. The case against discounting is even stronger, however, as Asheim and Buchholz (2003) demonstrate. They show that undiscounted utilitarianism is capable of obtaining any efficient and monotone sequence of intertemporal consumption levels as the unique solution within a reasonably wide class of feasible sets of allocations by suitable choice of a utility function. Thus, even in the absence of examples such as the one just presented, the argument in favour of discounting does not have much force. Our view is that, for the purpose of social evaluation, the well-being of future generations should not be discounted. If maximization of the ethically appropriate objective function requires the present generation to sacrifice most of its consumption for the benefit of others, then such an action can be considered supererogatory: desirable but beyond the call of duty. If these sacrifices are considered to be too demanding, we do not think it is a suitable response to give future generations a smaller weight in the social ordering. Instead, a sufficiently high level of well-being for the present generation can be guaranteed by imposing a floor on their utility as an additional constraint in the choice problem. This is a more natural and ethically attractive way of dealing with problems arising from supererogation than replacing an ethically appropriate social ordering with one that fails to treat generations impartially. See also Blackorby, Bossert and Donaldson (2000) for details. Cowen (1992) and Cowen and Parfit (1992) present a Paretian argument against discounting, and further discussions can be found in Broome (1992, pp. 92–108, 2004, pp. 126–8).

References

Aczél, J. (1966) *Lectures on Functional Equations and Their Applications* (New York: Academic Press, New York).

Asheim, G. and W. Buchholz (2003) 'The Malleability of Undiscounted Utilitarianism as a Criterion of Intergenerational Justice', *Economica*, vol. 70, pp. 405–22.

Blackorby, C., W. Bossert and D. Donaldson (1995) 'Intertemporal Population Ethics: Critical-level Utilitarian Principles', *Econometrica*, vol. 63, pp. 1303–20.

Blackorby, C., W. Bossert and D. Donaldson (1997) 'Birth-date Dependent Population Ethics: Critical-level Principles', *Journal of Economic Theory*, vol. 77, pp. 260–84.

Blackorby, C., W. Bossert and D. Donaldson (1999) 'Functional Equations and Population Ethics', *Aequationes Mathematicae*, vol. 58, pp. 272–84.

Blackorby, C., W. Bossert and D. Donaldson (2000) 'The Value of Limited Altruism', *Journal of Economic Theory*, vol. 95, pp. 37–70.

Blackorby, C., W. Bossert and D. Donaldson (2002) 'Utilitarianism and the Theory of Justice', in K. Arrow, A. Sen and K. Suzumura (eds), *Handbook of Social Choice and Welfare*, vol. 1 (Amsterdam: Elsevier), pp. 543–96.

Blackorby, C., W. Bossert and D. Donaldson (2005a) 'Multi-profile Welfarism: a Generalization', *Social Choice and Welfare*, vol. 24, pp. 253–67.

Blackorby, C., W. Bossert and D. Donaldson (2005b) *Population Issues in Social Choice Theory, Welfare Economics and Ethics* (Cambridge: Cambridge University Press).

Blackorby, C., D. Primont and R. Russell (1978) *Duality, Separability, and Functional Structure: Theory and Economic Applications* (Amsterdam: North-Holland).

Broome, J. (1992) *Counting the Cost of Global Warming* (Cambridge: White Horse).

Broome, J. (2004) *Weighing Lives*, (Oxford: Oxford University Press).

Cowen, T. (1992) 'Consequentialism implies a Zero Rate of Intergenerational Discount', in P. Laslett and J. Fishkin (eds), *Justice Between Age Groups and Generations* (New Haven: Yale University Press), pp. 162–8.

Cowen, T. and D. Parfit (1992) 'Against the Social Discount Rate', in P. Laslett and J. Fishkin (eds), *Justice Between Age Groups and Generations* (New Haven: Yale University Press), pp. 144–61.

d'Aspremont, C. (2007) 'Formal Welfarism and Intergenerational Equity', ch. 8 in this volume.

d'Aspremont, C. and L. Gevers (1977) 'Equity and the Informational Basis of Collective Choice', *Review of Economic Studies*, vol. 44, pp. 199–209.

Debreu, G. (1959) *Theory of Value: An Axiomatic Analysis of Economic Equilibrium* (New York: Wiley).

Debreu, G. (1960) 'Topological Methods in Cardinal Utility Theory', in K. Arrow, S. Karlin and P. Suppes (eds), *Mathematical Methods in the Social Sciences, 1959: Proceedings* (Stanford: Stanford University Press), pp. 16–26.
Fleming, M. (1952) 'A Cardinal Concept of Welfare', *Quarterly Journal of Economics*, vol. 66, pp. 366–84.
Goodin, R. (1991) 'Utility and the Good', in P. Singer (ed.), *A Companion to Ethics,* (Oxford: Blackwell), pp. 241–8.
Gorman, W. (1968) 'The Structure of Utility Functions', *Review of Economic Studies*, vol. 32, pp. 369–90.
Kelsey, D. (1987) 'The Role of Information in Social Welfare Judgements', *Oxford Economic Papers*, vol. 39, pp. 301–17.
Koopmans, T. (1960) 'Stationary Ordinal Utility and Impatience', *Econometrica*, vol. 28, pp. 287–309.
Maskin, E. (1978) 'A Theorem on Utilitarianism', *Review of Economic Studies*, vol. 45, pp. 93–6.
Roberts, K. (1980) 'Interpersonal Comparability and Social Choice Theory', *Review of Economic Studies*, vol. 47, pp. 421–39.
Suppes, P. (1966) 'Some Formal Models of Grading Principles', *Synthèse*, vol. 6, pp. 284–306.

10
Intergenerational Fairness*

Marc Fleurbaey

10.1 Introduction

This chapter studies the construction of a criterion for the ethical evaluation of allocations in an overlapping generations model. Each generation is composed of individuals who live two periods and may have heterogeneous intertemporal preferences. Their preferences are self-centred and are supposed to be a correct embodiment of their true personal interests. As a consequence, the criterion is required to satisfy the Pareto criterion. In addition, two basic fairness requirements are imposed on the criterion. The first result is then that the asymmetric part (strict preference) of the criterion must apply the infimum criterion (a variant of the maximin criterion suitable for infinite populations) to a particular money-metric utility representation of individual preferences. The choice of this particular utility measure is a consequence of the fairness requirements.

This result does not, however, fully characterize a complete ordering and the rest of the chapter studies how to define a complete ordering on the basis of the same requirements. An example of a criterion satisfying all the requirements is provided, and is shown to be the only one in a wide family of criteria.

This chapter is at the intersection of two literatures. It makes an extension to infinite populations of the analysis of fairness for finite populations made in Fleurbaey (2004, 2005). This part of the fairness literature (other references include, for example, Fleurbaey and Maniquet 2005, 2006; Maniquet and Sprumont 2004) relies on the Arrovian approach to social choice and studies

* This chapter is dedicated to Philippe Michel. It originates in a discussion with him and it would have been much better if it could have been co-written with him. It has benefitted from very helpful comments by T. Shinotsuka and from the reactions of participants at the International Economic Association Conference in Hakone. The hospitality of Nuffield College, Oxford, where this chapter was written, is gratefully acknowledged.

social orderings based on heterogeneous individual preferences, relying on axioms of efficiency and fairness. The kind of impossibility that Arrow (1951) famously obtained is avoided by weakening his axiom of Independence of Irrelevant Alternatives.

The other relevant part of the literature is formed by the numerous works analyzing the dilemma between efficiency and impartiality with infinite populations, in relation to the theory of optimal growth. The problem there is to find criteria that do not use the discounting method and nonetheless satisfy suitable versions of the Pareto principle. Among many references, one may cite Ramsey (1928), Koopmans (1965), Diamond (1965), Gale (1967), Brock (1970), Svensson (1980), Asheim (1991), Lauwers (1995, 1997a,b, 1998), Shinotsuka (1998), Fleurbaey and Michel (2003), Suzumura and Shinotsuka (2003), Basu and Mitra (2003), Sakai (2003, 2006), Asheim and Tungodden (2004a,b, 2005), Bossert, Sprumont and Suzumura (2004). Contrary to what is done here, most of this literature conveniently assumes that individual well-being is measured by an exogenously given utility function and that all individuals in a generation are identical and identically treated. Some results of this chapter are related to Fleurbaey and Michel (2003) because they make a similar use of the notion of ultrafilter in the construction of a criterion.

The results obtained here extend the positive results obtained in the theory of fair social choice, in the sense that a characterization of a particular measure of individual well-being and a precise description of the asymmetric part of the social criterion is obtained. They also extend or confirm some of the half-positive, half-negative results obtained in the theory of optimal growth, in the sense that while the existence of a complete ordering satisfying all the requirements is obtained, the definition of this ordering involves free ultrafilters and therefore eludes any fully explicit formulation.

The chapter is structured as follows. The next section introduces the model and the notations. The efficiency and fairness requirements imposed on the social criterion are presented in section 10.3. Sections 10.4 and 10.5 contain the results, and section 10.6 concludes. The mathematical proofs are in the appendix.

10.2 Model and notations

The set of real numbers is $\mathbb{R}$ (with $\mathbb{R}_+$ for non-negative and $\mathbb{R}_{++}$ for positive numbers), the set of natural integers is $\mathbb{N}$. Vector inequalities are denoted $\geq, >, \gg$ and set inclusions are denoted $\subset, \subsetneq$.

In the model considered in this chapter, there is only one physical good, but every individual lives two periods. Individual i's consumption is denoted $x_i = (c_i, d_i)$, with $c_i \geq 0$ the consumption in the first period, $d_i \geq 0$ the consumption in the second. Individual i has a preference relation R_i over $\mathbb{R}^2_+$. Let P_i, I_i denote the corresponding strict preference and indifference relations, respectively. Throughout we consider preference relations which are continuous, monotonic and convex.

The population N is fixed and contains an infinite succession of overlapping generations from $t = 0$ to infinity. For simplicity it is assumed that every generation contains the same number n of individuals. Let $t(i)$ denote the first period of life of individual $i \in N$.

There is a constant returns to scale technology (e.g. an international market with fixed price) which transforms one unit of good invested in one period into $1 + r$ units in the next period (with $r > 0$). Let $B(w, R_i)$ be the set of optimal bundles for i in the budget set with intertemporal wealth w:

$$B(w, R_i) = \left\{ x_i = \left(c_i, d_i\right) \in \mathbb{R}_+^2 \;\middle|\; \begin{array}{c} c_i + \frac{d_i}{1+r} \leq w \text{ and} \\ \forall (c, d) \in \mathbb{R}_+^2,\ c + \frac{d}{1+r} \leq w \Rightarrow x_i \, R_i \, (c, d) \end{array} \right\}.$$

Dually, let $E(x_i, R_i)$ denote the smallest intertemporal wealth which enables i to be as well-off as with x_i:

$$E(x_i, R_i) = \min \left\{ w \in \mathbb{R}_+ \mid \exists (c, d) \in \mathbb{R}_+^2,\ c + \frac{d}{1+r} \leq w,\ (c, d) \, R_i \, x_i \right\}.$$

This is a money-metric utility representation of R_i, as one has

$$E(x_i, R_i) \geq E(x_i', R_i) \Leftrightarrow x_i \, R_i \, x_i'.$$

An allocation is a list of all individuals' bundles: $x_N = (x_i)_{i \in N}$. A profile of preferences is similarly denoted $R_N = (R_i)_{i \in N}$. We focus on the subset of allocations with 'bounded resources', i.e. the subset defined as follows:

$$X = \left\{ x_N \in \left(\mathbb{R}_+^2\right)^N \mid \sum_{i \in N} \frac{1}{(1+r)^{t(i)}} \left[c_i + \frac{d_i}{1+r} \right] < +\infty \right\}.$$

This model is quite simple, in particular with respect to technology. It is nonetheless more complex than many versions of the overlapping generations model in that it allows every individual to have specific intertemporal preferences over bundles (c_i, d_i). An additional specific feature is that we do not have any information about utilities. Only ordinal non-comparable preferences are considered, as in the tradition of Arrovian social choice or in the theory of fairness as initiated by Kolm (1972), Varian (1974), Pazner and Schmeidler (1978). There are several possible justifications for this restriction on the information about individual well-being. First, one may pragmatically consider that information about comparable utilities is just not available, and not retrievable from individual behaviour in any reliable way. Second, one may consider that fairness is about providing resources to individuals, respecting their preferences, but is not about catering to idiosyncratic utility functions. For instance, a preference for consumption in period 1 rather than period 2 may justify letting the individual consume more in period 1, as a matter of respect for personal preference, or even simply of freedom of choice in one's budget set. In contrast, two individuals with identical preferences but different utility functions may be given the same resources and the one with a lower utility function is in practice held responsible for his lower level of utility. One may be afraid that this neglect of utilities is likely to be too harsh towards individuals whose lower utility function is due to a particular disability. The case of disabilities actually does not call for consideration of utilities, but simply for an extension of the scope of preferences to the relevant internal resources (see Fleurbaey and Maniquet, 1999, for a discussion and survey on such issues). Here it will be assumed that all individuals are equally endowed with internal resources and abilities, and that they can be held responsible for their utility functions, which will be considered to belong to the private sphere and will be totally disregarded.

A social ordering function (SOF) R defines an ordering (i.e. a complete and transitive binary relation) $R(R_N)$ over the set of allocations X for every profile R_N in a domain $\mathcal{D}$. Similarly, a social quasi-ordering function (SQOF) R defines a quasi-ordering (i.e. a reflexive and transitive binary relation) $R(R_N)$ over the set of allocations X for every profile R_N in a domain $\mathcal{D}$. Let $P(R_N)$ and $I(R_N)$ denote the strict preference and indifference counterparts of $R(R_N)$, respectively. The domain $\mathcal{D}$ considered in this chapter is the set of profiles R_N such that every R_i is continuous, monotonic and convex. The problem addressed in this chapter is the definition of a satisfactory SOF defined on this domain.

10.3 Axioms

The method adopted here follows Arrow's approach to social choice. A satisfactory SOF is one that complies with some basic conditions of efficiency and fairness. Although this axiomatic method is sometimes criticized as too abstract or too crude, one has to admit that there is no other game in town. There might be various ways of analyzing good and bad properties of SOFs, but ultimately they all boil down to defining requirements which can be formulated as axioms.

Efficiency concerns are represented here by a weak version of the Pareto principle. This principle is justified when individual preferences correctly reflect individuals' interests, and we assume it to be the case here. Such considerations are quite important in the context of intertemporal allocation, since intertemporal preferences, in real life, typically suffer from myopia. If individuals' ordinary preferences are not respectable because of such defects, then one should be able, at least in theory, to refer to the 'authentic' preferences that individuals would form if they were put in good conditions of information and deliberation.

Axiom 1 (Weak Pareto) $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X,$

$$[\forall i \in N, x_i P_i x'_i] \Rightarrow x_N P(R_N) x'_N.$$

The second condition is inspired by the Pigou–Dalton principle of transfer, and says that if an individual consumes more than another in the two periods of his life, then it would not be worse for the distribution if the 'rich' made a transfer, for the two periods of life, to the 'poor'. (This applies only to the case when they have identical preferences, in order to avoid a possible clash with the Pareto principle if this transfer condition were applied too widely. On this possible clash, see Fleurbaey and Trannoy, 2003).

Axiom 2 (Transfer) $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X, \forall i, j \in N, \forall \delta \in \mathbb{R}^2_{++}$, *if* $R_i = R_j, x_i = x'_i - \delta \gg x'_j + \delta = x_j$ and $\forall k \neq i, j, x_k = x'_k$, *then* $x_N R(R_N) x'_N$.

Like the basic Pigou–Dalton transfer principle usually applied to income distributions, this axiom, in isolation, is very weakly egalitarian. It is, for instance, satisfied by the overtaking criterion applied to the sequence of the 'market values' of generational consumptions and defined as follows:

$$x_N R(R_N) x'_N \Leftrightarrow \exists T^*, \forall T > T^*, \sum_{t=0}^{T} \sum_{\substack{i \in N \\ t(i)=t}} \left(c_i - c'_i + \frac{d_i - d'_i}{1+r} \right) \geq 0.$$

When the two individuals involved are not living at the same period, then the transfer mentioned in the axiom is either not feasible with the resources used in the allocation, or is wasteful. This is because transferring resources from one period to the next generates returns at the interest rate r. But the above axiom need not consider feasible transfers. It just makes a basic point, namely, that if such transfers were to be made, they would not worsen the social situation. This is a purely counterfactual observation. In this way, this axiom implies a symmetry of treatment between two individuals i and j with identical preferences but possibly different birth dates, in line with an ideal of impartiality between generations. Independently of when they happen to live, individuals are treated similarly and reducing inequality between them is considered reasonable. Through its egalitarian implications, this axiom has a definitely impartial content. Some comparison between this egalitarian axiom and a more limited anonymity requirement is made in the appendix (after the proof of Theorem 1).

The next condition considers how best to share a given amount of intertemporal wealth. It says that the egalitarian sharing of a given wealth, letting every individual choose in her own budget set defined by her equal wealth endowment, is the best allocation among all those that do not involve consuming more intertemporal wealth.

Axiom 3 (Equality) $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X, \forall w \in \mathbb{R}_{++}$, if $\forall i \in N$, $x_i \in B(w, R_i)$

$$\text{and} \qquad \sum_{i \in N} \frac{1}{(1+r)^{t(i)}} \left[c'_i + \frac{d'_i}{1+r} \right] \leq \left(1 + \frac{1}{r}\right) nw, \text{ then } x_N \, R(R_N) \, x'_N.$$

When a positive interest rate makes it easy to transfer resources to future generations, this appears as a substantial egalitarian condition, but it does not imply that for the sake of achieving equality it might be worth reducing the intertemporal wealth. Moreover it does not even require impartiality between generations since it is trivially satisfied by the simple criterion defined as follows:

$$x_N \, R(R_N) \, x'_N \Leftrightarrow \sum_{i \in N} \frac{1}{(1+r)^{t(i)}} \left[c_i + \frac{d_i}{1+r} \right] \geq \sum_{i \in N} \frac{1}{(1+r)^{t(i)}} \left[c'_i + \frac{d'_i}{1+r} \right].$$

This criterion favours earlier generations and does not even display any strict inequality aversion. This shows that the Equality axiom is not a strongly egalitarian axiom.

In order to make things even clearer, let us compare and illustrate the egalitarian import of the Transfer and Equality axioms in a simple two-period, one-good context. Assume that there are two individuals, living only one period each, at two successive dates, and consuming only one good. The set of feasible allocations is not symmetrical because of the interest rate and this creates a situation that is more favourable to the second-period individual. Figure 10.1 shows this configuration, where the slope of the possibility frontier is $-(1+r)$.

In this simple setting, the Transfer axiom says that any change of an allocation which moves the point that represents it in the figure toward the 45° line, perpendicularly (i.e. along arrows as in the figure), is a weak improvement. The Equality axiom says that the best allocation in the feasible set is depicted by the point at the intersection of the frontier with the 45° line. It is easy to see from this figure that none of these two axioms is strongly egalitarian, and that even the combination of the two can be satisfied by very weakly egalitarian social orderings. Figure 10.2 shows the social indifference curves of such an ordering.

The next and last condition is related to Arrow's Independence of Irrelevant Alternatives. It is logically weaker because it allows the ranking of two allocations to depend not only on how they are ranked by individual preferences (as in Arrow's condition), but also on the indifference curves of individuals at the considered bundles. Taking account of indifference curves is a common practice in the theory of fairness, in cost–benefit analysis, etc.

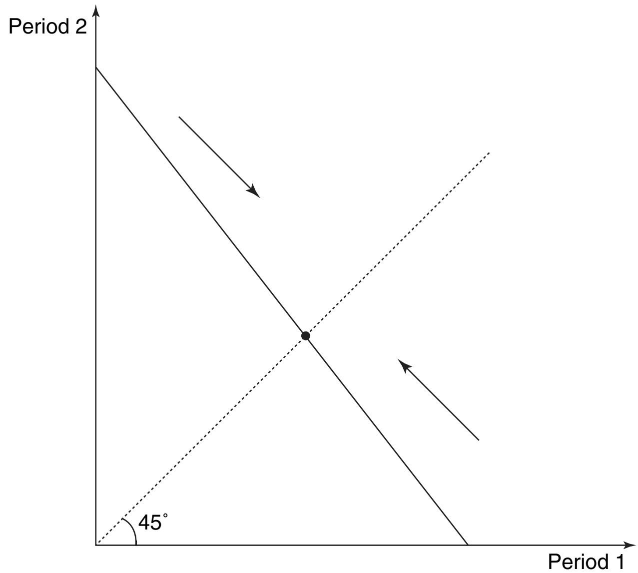

Figure 10.1 Illustration of the transfer axiom

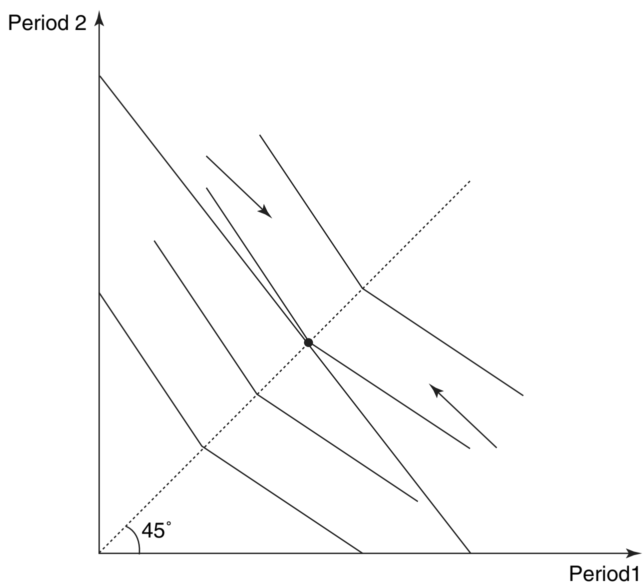

Figure 10.2 Illustration of the equality axiom

It can indeed be argued that this additional information is ethically relevant in order to thoroughly evaluate and compare individual situations. For instance, when individual i prefers x_i to x'_i whereas individual j prefers x'_j to x_j, it may be quite worthwhile to be left to know, additionally, that both individuals consider that x_i is a better bundle than x_j whereas they would both put x'_i and x'_j on a par. More discussion of the independence conditions can be found in Fleurbaey and Maniquet (2006) and Fleurbaey, Suzumura and Tadenuma (2003). Typically, an independence condition of this kind can be motivated by the desire to make the SOF informationally parsimonious or computationally simple, by an ethical principle of responsibility saying that the evaluation of allocations should not be too sensitive to individual preferences (because individuals must assume responsibility for their preferences, to some extent), or by implementation concerns.

Axiom 4 (Independence) $\forall R_N, R'_N \in \mathcal{D}, \forall x_N, x'_N \in X$,
if $\forall i \in N, \forall y \in \mathbb{R}^2_+, y\, I_i\, x_i \Leftrightarrow y\, I'_i\, x_i$ and $y\, I_i\, x'_i \Leftrightarrow y\, I'_i\, x'_i$,
then $x_N\, R(R_N)\, x'_N \Leftrightarrow x_N\, R(R'_N)\, x'_N$.

In summary, we have four requirements, one reflecting efficiency concerns (Weak Pareto), two reflecting fairness concerns (Transfer and Equality), and one reflecting various possible concerns of informational parsimony or responsibility (Independence). Our aim is to find and describe SOFs and SQOFs that satisfy the four axioms.

10.4 Social ordering functions

The first result below gives some indications about two things. First, it says that the asymmetric part of the SOF must involve the infimum criterion, which is a quite strongly egalitarian conclusion. Second, it says that, in the application of the infimum criterion, individual well-being must be measured by the money-metric utility function $E(x_i, R_i)$. This double characterization of social preferences and individual well-being is typical of results obtained in the theory of fair social choice.

Theorem 1 *Let R satisfy the four axioms. Then the following is true:* $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X$,

$$\inf_{i\in N} E(x_i, R_i) > \inf_{i\in N} E(x'_i, R_i) \Rightarrow x_N\, P(R_N)\, x'_N.$$

All the proofs are in the appendix at the end of this chapter. This result extends similar results in Fleurbaey (2004, 2005) in two ways. First, it applies to infinite populations, which implies a complication in the analysis of allocations since the Transfer axiom can be used only in a finite sequence of transfers. Second, it involves budget sets and a money-metric utility function whereas the quoted works consider a simpler metric.

An additional difference comes from the technological asymmetry between different generations in the current framework. Later generations benefit from the fact that transferring resources from earlier generations to them generates a surplus (the interest rate r). As a consequence, a standard utilitarian-type objective with no discounting (such as the overtaking criterion applied to $U(x_i)$ for some concave utility function U) would typically select a consumption path with positive growth. It appears that the axioms Transfer and Equality both contribute to counteracting this bias in favour of later generations, but the Transfer axiom is more decisive in generating the infimum result. In particular, the conclusion of the theorem would still hold if Equality were replaced by an axiom saying that there is a ranking R^* over $\mathbb{R}_+^N$ such that for all $R_N \in \mathcal{D}$, all $x_N, x'_N \in X$, all $w_N, w'_N \in \mathbb{R}_+^N$, if for all $i \in N$, $x_i \in B(w_i, R_i)$ and $x'_i \in B(w'_i, R_i)$, then

$$x_N \, R(R_N) \, x'_N \Leftrightarrow w_N \, R^* \, w'_N.$$

This axiom says nothing about the distribution and in particular has no egalitarian implication whatsoever. It simply says that when allocations are generated by individuals making free choices in their budget sets, one can simply rank the distribution of wealth and there is no need to examine the composition of personal consumption bundles in more detail. This is a quite reasonable idea in this context since preferences are assumed to be fully respectable.

The extension to infinite populations entails yet another difference. In the study of finite populations, one typically obtains a similar result referring to the *maximin* criterion. Although this is not a full characterization of the maximin criterion (applied to the suitable metric of individual well-being), because only the asymmetric part is described in the result, there is some sense in which the maximin criterion appears as the natural option in this case, if one looks for a SOF satisfying all the conditions of the theorem (the leximin is another obvious option if one thinks of strenghtening the Pareto requirement). Here, however, things are quite different. The infimum criterion does not satisfy the Weak Pareto axiom, and no obvious example of a SOF satisfying all the four axioms is readily available. One could even be afraid that the four axioms are incompatible, so that the theorem would simply be an absurd result due to the fact that a non-existing object can be proved to satisfy any property! We therefore have to look for such an example.

The next result provides one, and relies on the notion of ultrafilter in a similar fashion as in Fleurbaey and Michel (2003). Let $\mathcal{F}$ be a set of subsets of $\mathbb{N}$ satisfying the following properties:

(i) $\emptyset \notin \mathcal{F}$;
(ii) $\forall A, B \in \mathcal{F}, A \cap B \in \mathcal{F}$;
(iii) $\forall A \in \mathcal{F}, \forall B \subset \mathbb{N}, A \subset B \Rightarrow B \in \mathcal{F}$;

(iv) $\forall A \subset \mathbb{N}, A \in \mathcal{F}$ or $\mathbb{N} \setminus A \in \mathcal{F}$;
(v) $\forall A \in \mathcal{F}, \#A = \infty$.

These properties make $\mathcal{F}$ a free ultrafilter. The existence of free ultrafilters is a well-known mathematical result (see e.g. Bourbaki, 1961). A simple fact is useful to bear in mind: It follows from (iv) and (v) that every free ultrafilter contains all subsets of the kind

$$A = \{T, T+1, \ldots\}.$$

Ultrafilters have been used in the theory of social choice in various ways but they have generally been defined on the set of individuals (see Kirman and Sondermann, 1972, Monjardet, 1983, Lauwers and Van Liedekerke, 1995), whereas here, as in Fleurbaey and Michel (2003), they are defined on the set of dates.

Theorem 2 *Let $\mathcal{F}$ be a free ultrafilter on $\mathbb{N}$. Let R be defined as follows:* $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X$,

$$x_N \, R(R_N) \, x'_N \;\Leftrightarrow$$
$$\exists A \;\in\; \mathcal{F}, \forall T \in A, \; \min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(x'_i, R_i).$$

Then R satisfies the four axioms.

This positive result is not totally satisfactory. The existence of an ultrafilter is usually proved not by explicit construction but with the help of the Axiom of Choice or similar premises.[1] Therefore the above example is not fully explicit. If one is given two arbitrary allocations $x_N, x'_N \in X$, it may be impossible to decide which one is best for R. One should therefore look for more explicit examples. This is the object of the next result, which says that within the family of SOFs which rely on the maximin criterion applied to a subset of horizons, there is no hope to satisfy the axioms of Theorem 1 without relying on the notion of free ultrafilter.

The next result requires an additional definition. Let $\mathcal{E}$ be a set of subsets of $\mathbb{N}$ satisfying (i), (v) and the following modifications of (ii) and (iv):

(ii′) $\forall A, B \in \mathcal{E}, \exists C \in \mathcal{E}, C \subset A \cap B$;
(iv′) $\forall A \subset \mathbb{N}, \exists B \in \mathcal{E}, B \subset A$ or $B \subset \mathbb{N} \setminus A$.

Then $\mathcal{E}$ is said to be the base of a free ultrafilter, because the set $\mathcal{F}$, defined by: $\forall A \subset \mathbb{N}$,

$$A \in \mathcal{F} \Leftrightarrow \exists B \in \mathcal{E}, \; B \subset A,$$

is a free ultrafilter.

Theorem 3 *Let $\mathcal{E}$ be a set of subsets of $\mathbb{N}$ and let R be defined as follows:* $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X,$

$$x_N\, R(R_N)\, x'_N \Leftrightarrow$$
$$\exists A \in \mathcal{E}, \forall T \in A, \min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(x'_i, R_i).$$

Then R satisfies the four axioms if and only if $\mathcal{E}$ is the base of a free ultrafilter.

This result is not a general proof that there is no constructible example of a SOF satisfying the four axioms of Theorem 1. But it shows that in the family of SOFs involving the maximin criterion applied to subsets of horizons, there is indeed an impossibility to find anything different from the example of Theorem 2.

The proof of Theorem 3 (in the Appendix) shows that the key properties in generating the free ultrafilter result are completeness, transitivity of the SOF, Weak Pareto and Transfer. In this sense, this result confirms and illustrates again the well-known difficulty of finding complete orderings respecting the Pareto principle and basic impartiality requirements for an infinite population.

10.5 Social quasi-ordering functions

This difficulty underlies the interest for quasi-orderings (i.e. partial rankings) such as the overtaking criterion in the theory of optimal growth. Relaxing the requirement of completeness may indeed appear worthwhile if it can be compensated by the possibility of an explicit definition of the criterion.

It turns out that Theorem 1 above applies to SQOFs exactly as it applies to SOFs, and this can be seen by the fact that its proof is constructed in a way that does not appeal to completeness.[2] Therefore the axioms considered here are as constraining over the asymmetric part of a SQOF as they are for a SOF.

It remains to find constructible examples of SQOFs satisfying the axioms. The following theorem provides one such example, which is again inspired by the overtaking criterion.

Theorem 4 *Let R be defined as follows:* $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X,$

$$x_N\, R(R_N)\, x'_N \Leftrightarrow$$
$$\exists T^* \in \mathbb{N}, \forall T \geq T^*, \min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(x'_i, R_i),$$

This SQOF satisfies the four axioms.

It is customary to also consider a more continuous version of the overtaking criterion, defined as

$$x_N\, R(R_N)\, x'_N \Leftrightarrow \liminf_{T\to\infty} \left[\min_{\substack{i\in N\\ t(i)\leq T}} E(x_i, R_i) - \min_{\substack{i\in N\\ t(i)\leq T}} E(x'_i, R_i) \right] \geq 0.$$

But the sequences

$$a_T = \min_{\substack{i\in N\\ t(i)\leq T}} E(x_i, R_i),\ \ b_T = \min_{\substack{i\in N\\ t(i)\leq T}} E(x'_i, R_i)$$

are non-increasing and bounded below, so that their limits are well defined and

$$\lim_{T\to\infty} a_T - \lim_{T\to\infty} b_T = \lim_{T\to\infty} (a_T - b_T) = \liminf_{T\to\infty} (a_T - b_T) = \limsup_{T\to\infty} (a_T - b_T).$$

Moreover one has

$$\lim_{T\to\infty} a_T = \inf_{i\in N} E(x_i, R_i).$$

Therefore this overtaking criterion is actually equivalent to the infimum criterion (a SOF, not just a SQOF), which does not satisfy Weak Pareto.

10.6 Conclusion

This set of results suggests a conjecture similar to one made in a different setting in Fleurbaey and Michel (2003). There might be no constructible SOF satisfying the four axioms considered in this chapter. Moreover, as mentioned above, the problem really involves only two of the axioms: Weak Pareto and Transfer. The other two axioms do play an important role in obtaining the positive result of Theorem 1 but are not essential to the constructibility issue.

How important is this problem? If one restricts attention to non-decreasing sequences of consumption (as evaluated by the $E(x_i, R_i)$ function), the problem vanishes because the infimum criterion is then identical to the maximin, which in this case satisfies the four conditions. Sustainability, in this sense, comes to the rescue of optimal growth. Moreover, in applications, generally, the infimum criterion or partial criteria such as those presented in Theorem 4 do have a substantial discriminatory power and point to a narrow subset of optimal paths.

One may then consider that the positive side of the above results is more important than the negative side. The conclusion of Theorem 1, however, may raise objections. The strongly egalitarian flavour of the infimum criterion is a systematic growth killer. Theorem 1 may at least be useful in forcing the opponents to the infimum to explain what conditions of the theorem they reject, and on what basis. The maximin–leximin–infimum family of criteria is, however, far from being ignored in the literature on optimal growth (see e.g. Arrow, 1973; Dasgupta, 1974; Lauwers, 1997b; Roemer and Veneziani, 2003; Asheim and Tungodden, 2004b). This chapter brings some additional arguments in its favour. Even if such egalitarian conclusions were not ultimately retained, this chapter might help in clarifying the ethical underpinnings of social criteria for growth.

One important simplification of this analysis comes from the linear technology (a fixed interest rate r). This assumption makes it possible to focus on $E(x_i, R_i)$, the computation of which depends on r, as the relevant metric of well-being. With a more complex technology, the determination of the ethically appropriate metric of well-being would be less obvious. One must not, however, be mistaken about the role of the technology in this analysis. All the results of this chapter would go through with any kind of technology, because nothing in them formally depends on the technology. The Equality axiom is the only requirement which relies on r, but it may be retained even when the actual interest rate may vary depending on the path, if one considers that the parameter r that appears in Equality is just a benchmark value serving to formulate an egalitarian judgment in a simple case (the case when the interest rate happens to be fixed at the value r). The technological hypothesis made here only serves to make r an obviously salient value and to make Equality a plausible condition. With a more flexible technology, the choice of r as a benchmark value in the Equality axiom would sound a little arbitrary. One may then consider that another kind of Equality axiom would be needed, leading to a different metric of well-being. This particular extension of the present analysis is left for future research.

Appendix: Proofs

We need the following notations for upper and lower contour sets:

$$U(x_i, R_i) = \left\{q \in \mathbb{R}^2_+ \mid q\, R_i\, x_i\right\},$$
$$L(x_i, R_i) = \left\{q \in \mathbb{R}^2_+ \mid x_i\, R_i\, q\right\}.$$

Proof of Theorem 1: The proof relies on the following lemmas.

Lemma 1 *Let R satisfy Weak Pareto, Transfer and Independence. Then the following is true:* $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X, \forall i, j \in N,$
if $R_i = R_j, x'_i\, P_i\, x_i\, P_i\, x_j\, P_j\, x'_j$ *and* $\forall k \neq i, j, x_k\, P_k\, x'_k,$
then $x_N\, P(R_N)\, x'_N.$

Proof: See Lemma 1 in Fleurbaey (2004). □

Lemma 2 *Let R satisfy Weak Pareto, Transfer and Independence. Then the following is true:* $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X$,
if $\exists G \subset N$, $\#G < \infty$, *such that:*
$\forall i \in G$, $x'_i \, P_i \, x_i$ *and* $\exists j \in N \setminus G$, $x_j \, P_j \, x'_j$ *and* $U(x_i, R_i) \cap L(x_j, R_j) = \emptyset$, *and* $\forall i \in N \setminus G$, $x_i \, P_i \, x'_i$, *then* $x_N \, P(R_N) \, x'_N$.

Proof: The proof relies on Lemma 1 and on the fact that if $x'_i \, P_i \, x_i$, $x_j \, P_j \, x'_j$ and $U(x_i, R_i) \cap L(x_j, R_j) = \emptyset$, then there exists R'_0 such that

$$\begin{aligned} U(x_i, R_0) &= U(x_i, R_i),\ U(x'_i, R_0) = U(x'_i, R_i), \\ U(x_j, R_0) &= U(x_j, R_j),\ U(x'_j, R_0) = U(x'_j, R_j), \\ & x'_i \, P_0 \, x_i \, P_0 \, x_j \, P_0 \, x'_j. \end{aligned}$$

For a detailed proof of a similar statement, see the proof of Lemma 2 in Fleurbaey (2004). □

Let $R_N \in \mathcal{D}$ and $x_N, x'_N \in X$ be such that

$$\inf_{i \in N} E(x_i, R_i) > \inf_{i \in N} E(x'_i, R_i).$$

Let $i_0 \in N$ be such that $\inf_{i \in N} E(x_i, R_i) > E(x'_{i_0}, R_{i_0})$. Pick $w, w^-, w^+ \in \mathbb{R}_{++}$, $\varepsilon \in \mathbb{R}_{++}$ and $T \in \mathbb{N}$ such that

$$\inf_{i \in N} E(x_i, R_i) > w^+ > w > w^- > E(x'_{i_0}, R_{i_0})$$

and

$$\frac{w^-}{(1+r)^{t(i_0)}} + \sum_{\substack{i \in N \setminus \{i_0\} \\ t(i) \le T}} \frac{w^+}{(1+r)^{t(i)}} + \sum_{\substack{i \in N \setminus \{i_0\} \\ t(i) > T}} \frac{1}{(1+r)^{t(i)}} \left[c'_i + \varepsilon + \frac{d'_i + \varepsilon}{1+r} \right] < \left(1 + \frac{1}{r}\right) nw.$$

This is made possible by

$$\sum_{i \in N} \frac{1}{(1+r)^{t(i)}} \left[c'_i + \frac{d'_i}{1+r} \right] < +\infty,$$

which implies

$$\lim_{T \to \infty} \sum_{\substack{i \in N \\ t(i) > T}} \frac{1}{(1+r)^{t(i)}} \left[c'_i + \varepsilon + \frac{d'_i + \varepsilon}{1+r} \right] = 0.$$

Let $x_N^- \in X$ be such that for all $i \in N$, $x_i^- \in B(w, R_i)$. For every $i \in N \backslash \{i_0\}$ such that $t(i) \leq T$, one can construct R_i' and $y_i, y_i', y_i'' \in \mathbb{R}_+^2$ such that $y_i'' \in B(w^+, R_i')$, $x_i^- \in B(w, R_i')$,

$$y_i \, P_i \, x_i, \; y_i \, P_i \, x_i', \; y_i' P_i \, y_i$$

and

$$\begin{aligned} U(y_i'', R_i') \cap L(x_{i_0}^-, R_{i_0}) &= \emptyset, \\ U(y_i', R_i') \cap L(y_i, R_i) &= \emptyset, \\ U(x_i, R_i) \cap L(x_i^-, R_i') &= \emptyset. \end{aligned}$$

For every $i \in N \backslash \{i_0\}$ such that $t(i) > T$, let $R_i' = R_i$ and $y_i, y_i', y_i'' \in \mathbb{R}_+^2$ be such that

$$x_i' + (\varepsilon, \varepsilon) \gg y_i'' \gg y_i' \gg y_i \gg x_i'.$$

And let $R_{i_0}' = R_{i_0}$ and $y_{i_0}, y_{i_0}', y_{i_0}'' \in \mathbb{R}_+^2$ be such that $y_{i_0}'' \in B(w^-, R_{i_0})$ and

$$y_{i_0}'' \, P_{i_0} \, y_{i_0}' \, P_{i_0} \, y_{i_0} \, P_{i_0} \, x_{i_0}'.$$

Let $R_N'' \in \mathcal{D}$ be such that: for every $i \in N \backslash \{i_0\}$ such that $t(i) \leq T$,

$$\begin{aligned} U(x_i, R_i'') &= U(x_i, R_i), U(y_i, R_i'') = U(y_i, R_i), \\ U(y_i', R_i'') &= U(y_i', R_i'), \; U(x_i^-, R_i'') = U(x_i^-, R_i'); \end{aligned}$$

for all other $i \in N$, $R_i'' = R_i$.

By Equality, $x_N^- \, R(R_N') \, y_N''$, since, denoting $y_i'' = (c_i'', d_i'')$,

$$\begin{aligned} &\sum_{i \in N} \frac{1}{(1+r)^{t(i)}} \left[c_i'' + \frac{d_i''}{1+r} \right] \\ &< \frac{w^-}{(1+r)^{t(i_0)}} + \sum_{\substack{i \in N \backslash \{i_0\} \\ t(i) \leq T}} \frac{w^+}{(1+r)^{t(i)}} + \sum_{\substack{i \in N \backslash \{i_0\} \\ t(i) > T}} \frac{1}{(1+r)^{t(i)}} \left[c_i' + \varepsilon + \frac{d_i' + \varepsilon}{1+r} \right] \\ &< \left(1 + \frac{1}{r}\right) nw. \end{aligned}$$

By Lemma 2, $y_N'' \, P(R_N') \, y_N'$, so that by transitivity, $x_N^- \, P(R_N') \, y_N'$. By Independence, $x_N^- \, P(R_N'') \, y_N'$.

By Weak Pareto, $x_N \, P(R_N'') \, x_N^-$ and $y_N' \, P(R_N'') \, y_N$ so that by transitivity, $x_N \, P(R_N'') \, y_N$. By Independence, $x_N \, P(R_N) \, y_N$. By Weak Pareto, $y_N \, P(R_N) \, x_N'$ so that by transitivity, $x_N \, P(R_N) \, x_N'$.

We finally check that the conditions of the theorem all play a role in the conclusion.

If Weak Pareto is dropped, the SOF displaying total indifference between all allocations satisfies the three other conditions.

If Transfer is dropped, let R be defined by:

$$x_N\, R(R_N)\, x'_N \Leftrightarrow \sum_{i\in N} \frac{1}{(1+r)^{t(i)}} E(x_i, R_i) \geq \sum_{i\in N} \frac{1}{(1+r)^{t(i)}} E(x'_i, R_i).$$

It is well defined over X and satisfies the three other axioms. (Notice that it relies on discounting.)

If Equality is dropped, one may replace $E(x_i, R_i)$ with any other metric $\widehat{E}(x_i, R_i)$ which correctly represents individual preferences.

If Independence is dropped, let R coincide with the example of Theorem 2 for all profiles except when R_i is linear for all i. Then R coincides with the SOF defined in the remark below. □

Remark. One might think that Transfer could be weakened into an Anonymity condition of the following kind:

Axiom 5 (Anonymity) $\forall R_N \in \mathcal{D}, \forall x_N, x'_N \in X, \forall i, j \in N,$
if $R_i = R_j$, $x'_i = x_j$, $x'_j = x_i$ *and* $\forall k \neq i, j$, $x_k = x'_k$,
then $x_N\, I(R_N)\, x'_N$.

Here is, however, a SOF which satisfies Weak Pareto, Anonymity, Equality and Independence and is quite different from the infimum criterion. For any $t \in \mathbb{N}$, any $x_N \in X$, let $S_t^T(x_N)$ be equal to the sum of $E(x_i, R_i)$ of the individuals who are ranked (according to the increasing order of $E(x_i, R_i)$) at the $(t-1)n+1$ through tn-th ranks, in the population of $i \in N$ such that $t(i) \leq T$. More precisely, let $(E_k^T)_{k=1,\ldots,nT}$ be the rearranged vector of $(E(x_i, R_i))_{\substack{i\in N\\ t(i)\leq T}}$ by increasing order. One then has

$$S_t^T(x_N) = \sum_{k=(t-1)n+1}^{tn} E_k^T$$

We define R as follows: $x_N\, R(R_N)\, x'_N$ iff

$$\exists A \in \mathcal{F}, \forall T \in A,\ \sum_{t\leq T} \frac{1}{(1+r)^t} S_t^T(x_N) \geq \sum_{t\leq T} \frac{1}{(1+r)^t} S_t^T(x'_N),$$

where $\mathcal{F}$ is a free ultrafilter on $\mathbb{N}$.

Proof of Theorem 2: *Weak Pareto*: Let $x_N, x'_N \in X$ be such that $x_i P_i x'_i$ for all $i \in N$. Then for all $A \subset \mathbb{N}$, $A \neq \emptyset$, one has

$$\forall T \in A, \min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) > \min_{\substack{i \in N \\ t(i) \leq T}} E(x'_i, R_i).$$

Transfer: Let $x_N, x'_N \in X$ and $i, j \in N$, $\delta \in \mathbb{R}^2_{++}$ be such that $R_i = R_j$, $x_i = x'_i - \delta \gg x'_j + \delta = x_j$ and $\forall k \neq i, j$, $x_k = x'_k$. This entails

$$\min_{k \in \{i,j\}} E(x_k, R_k) > \min_{k \in \{i,j\}} E(x'_k, R_k),$$

and for all $G \subset N$ such that $i, j \in G$,

$$\min_{k \in G} E(x_k, R_k) \geq \min_{k \in G} E(x'_k, R_k).$$

Let $T_0 \in \mathbb{N}$ be such that $t(i), t(j) \leq T_0$, and let

$$A = \{T_0, T_0 + 1, \ldots\}.$$

Necessarily $A \in \mathcal{F}$ (by (iv) and (v)) and

$$\forall T \in A, \min_{\substack{k \in N \\ t(k) \leq T}} E(x_k, R_k) \geq \min_{\substack{k \in N \\ t(k) \leq T}} E(x'_k, R_k).$$

Equality: Let $x_N, x'_N \in X$ and $w \in \mathbb{R}_{++}$ be such that $\forall i \in N$, $x_i \in B(w, R_i)$ and

$$\sum_{i \in N} \frac{1}{(1+r)^{t(i)}} \left[c'_i + \frac{d'_i}{1+r} \right] \leq \left(1 + \frac{1}{r} \right) nw.$$

By definition, for all $i \in N$, $E(x_i, R_i) = w$, and there is $i_0 \in N$, $E(x'_{i_0}, R_{i_0}) \leq w$. Let $T_0 \in \mathbb{N}$ be such that $t(i_0) \leq T_0$, and let

$$A = \{T_0, T_0 + 1, \ldots\}.$$

Necessarily $A \in \mathcal{F}$ and

$$\forall T \in A, \min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) = w \geq E(x'_{i_0}, R_{i_0}) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(x'_i, R_i).$$

Independence: This follows directly from the fact that the value of $E(x_i, R_i)$ depends only on $U(x_i, R_i)$. □

Proof of Theorem 3: The 'if' part is proved like Theorem 2. The only change is that the subset

$$A = \{T_0, T_0 + 1, \ldots\}$$

need not belong to $\mathcal{E}$ but by (iv') and (v) necessarily there is $B \, in \, \mathcal{E}$ such that $B \subset A$.

We now turn to the 'only if' part. We have to show that $\mathcal{E}$ must satisfy properties (i),(ii'),(iv') and (v).

(i) Suppose that $\emptyset \in \mathcal{E}$. Since for any $x_N, x'_N \in X$ one always has

$$\forall T \in \emptyset, \ \min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(x'_i, R_i)$$

then R would display universal indifference, in violation of Weak Pareto. Therefore $\emptyset \notin \mathcal{E}$.

(ii′) Let $A, B \in \mathcal{E}$. We construct $x_N, y_N, z_N \in X$ and $R_N \in \mathcal{D}$ such that for all i, j, $t(i) = t(j)$, one has $E(x_i, R_i) = E(x_j, R_j)$ and similarly for y_N, z_N. Let $E_t(x_N)$ denote $E(x_i, R_i)$ for any $i \in N$ such that $t(i) = t$. The allocations are constructed so as to have (for some $\alpha < 1$):

$$E_t(x_N) = \alpha^t \text{ for } t \in \mathbb{N};$$

$$E_t(y_N) = \begin{cases} \alpha^t \text{ for } t \in A, \\ 1 \text{ for } t \in \mathbb{N} \backslash A; \end{cases}$$

$$E_t(z_N) = \begin{cases} \min_{s \leq t} E_s(y_N) \text{ for } t \in B, \\ 1 \text{ for } t \in \mathbb{N} \backslash B. \end{cases}$$

One then has

$$\min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(y_i, R_i) \Leftrightarrow T \in A,$$

$$\min_{\substack{i \in N \\ t(i) \leq T}} E(y_i, R_i) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(z_i, R_i) \Leftrightarrow T \in B,$$

$$\min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(z_i, R_i) \Leftrightarrow T \in A \cap B.$$

The first two equivalences entail $x_N \, R(R_N) \, y_N$ and $y_N \, R(R_N) \, z_N$. By transitivity, $x_N \, R(R_N) \, z_N$. Therefore there exists $C \in \mathcal{E}$ such that

$$\forall T \in C, \ \min_{\substack{i \in N \\ t(i) \leq T}} E(x_i, R_i) \geq \min_{\substack{i \in N \\ t(i) \leq T}} E(z_i, R_i).$$

In view of the third equivalence above, this implies $C \subset A \cap B$.

(iv′) Let $A \subset \mathbb{N}$. We construct $x_N, y_N \in X$ and $R_N \in \mathcal{D}$ as in the previous paragraph. We then have

$$\min_{\substack{i\in N\\ t(i)\le T}} E(x_i,R_i) \ge \min_{\substack{i\in N\\ t(i)\le T}} E(y_i,R_i) \Leftrightarrow T \in A.$$

Since R is complete, either $x_N\, R(R_N)\, y_N$ or $y_N\, R(R_N)\, x_N$. In the former case, there exists $B \in \mathcal{E}$ such that

$$\forall T \in B,\ \min_{\substack{i\in N\\ t(i)\le T}} E(x_i,R_i) \ge \min_{\substack{i\in N\\ t(i)\le T}} E(y_i,R_i).$$

This implies $B \subset A$. In the latter case, there exists $B \in \mathcal{E}$ such that

$$\forall T \in B,\ \min_{\substack{i\in N\\ t(i)\le T}} E(y_i,R_i) \ge \min_{\substack{i\in N\\ t(i)\le T}} E(x_i,R_i).$$

This implies $B \subset \mathbb{N}\backslash A$.

(v) Let $A \in \mathcal{E}$ be such that $\#A < \infty$. Let $T_A = \max A$. Let $R_N \in \mathcal{D}$ be such that for all $i, j \in N$, $R_i = R_j$. We construct $x_N, y_N \in X$ such that for all $i, j \in N$, $E(x_i, R_i) = E(x_j, R_j)$ whenever $t(i) = t(j)$, and similarly for y_N. We can therefore use the same notation $E_t(x_N)$ as above. The allocations are constructed so as to have:

$$E_t(x_N) = \begin{cases} 10 \text{ for } t \le T_A, \\ 6 \text{ for } t > T_A; \end{cases}$$

$$E_t(y_N) = \begin{cases} 12 \text{ for } t \le T_A, \\ 4 \text{ for } t > T_A; \end{cases}$$

One then has

$$\forall T \in A,\ \min_{\substack{i\in N\\ t(i)\le T}} E(y_i,R_i) \ge \min_{\substack{i\in N\\ t(i)\le T}} E(x_i,R_i),$$

implying $y_N\, R(R_N)\, x_N$. But by Lemma 1, one must have $x_N\, P(R_N)\, y_N$, a contradiction. Therefore necessarily $\#A = \infty$ for all $A \in \mathcal{E}$. □

Proof of Theorem 4: It mimics the proof of Theorem 2, which in particular refers to subsets of the kind $A = \{T_0, T_0 + 1, \ldots\}$. □

Notes

1. It would have been possible to state the *existence* of a social ordering function satisfying the axioms, by making use of Szpilrajn's Lemma, as in Svensson (1980). This could be done by first exhibiting a SQOF (as in Theorem 4 below) and then extending it into a SOF. The above theorem is more explicit about the structure of the example.
2. That is another difference with the proofs one finds in Fleurbaey (2004, 2005).

References

Arrow, K. J. (1951) *Social Choice and Individual Values* (New York: Wiley).

Arrow, K. J. (1973) 'Rawls's Principle of Just Saving', *Swedish Journal of Economics*, vol. 75, pp. 323–35.

Asheim, G. B. (1991) 'Unjust Intergenerational Allocations', *Journal of Economic Theory*, vol. 54, pp. 350–71.

Asheim, G. B. and Tungodden, B. (2004a) 'Do Koopmans Postulates Lead to Discounted Utilitarianism?', Discussion paper 32/04, Norwegian School of Economics and Business Administration.

Asheim, G. B. and Tungodden, B. (2004b) 'Resolving Distributional Conflicts between Generations', *Economic Theory*, vol. 24, pp. 221–30.

Asheim, G. B. and Tungodden, B. (2005) 'A New Equity Condition for Infinite Utility Streams and the Possibility of Being Paretian', mimeo, University of Oslo.

Basu, K. and Mitra, T. (2003) 'Aggregating Infinite Utility Streams with Intergenerational Equity: The Impossibility of Being Paretian', *Econometrica*, vol. 71, pp. 1557–63.

Bossert, W., Sprumont, Y. and Suzumura, K. (2004) 'The Possibility of Ordering Infinite Utility Streams', Cahiers CIREQ 12-2004.

Bourbaki, N. (1961) *Topologie Générale* (Paris: Hermann).

Brock, W. A. (1970) 'On Existence of Weakly Maximal Programmes in a Multisector Economy', *Review of Economic Studies*, vol. 37, pp. 275–80.

Dasgupta, P. S. (1974) 'On Some Alternative Criteria for Justice between Generations', *Journal of Public Economics*, vol. 3, pp. 415–23.

Diamond, P. A. (1965) 'The Evaluation of Infinite Utility Streams', *Econometrica*, vol. 33, pp. 170–7.

Fleurbaey, M. (2004) 'Two Criteria for Social Decisions', Nuffield College Economics Working Paper 2004-W27.

Fleurbaey, M. (2005) 'The Pazner–Schmeidler Ordering: A Defense', *Review of Economic Design*, vol. 9, pp. 145–6.

Fleurbaey, M. and F. Maniquet (2006) 'Utilitarianism versus Fairness in Welfare Economics', in M. Salles and J. A. Weymark (eds), *Justice, Political Liberalism and Utilitarianism: Themes from Harsanyi and Rawls* (Cambridge: Cambridge University Press).

Fleurbaey, M. and F. Maniquet (1999) 'Compensation and Responsibility', in K. J. Arrow, A. K. Sen, K. Suzumura (eds), *Handbook of Social Choice and Welfare*, vol. 2 (Amsterdam: North-Holland).

Fleurbaey, M. and F. Maniquet (2005) 'Fair Social Orderings when Agents Have Unequal Production Skills', *Social Choice and Welfare*, vol. 24, pp. 93–127.

Fleurbaey, M. and P. Michel (2003) 'Intertemporal Equity and the Extension of the Ramsey Criterion', *Journal of Mathematical Economics*, vol. 39, pp. 777–802.

Fleurbaey, M., K. Suzumura and K. Tadenuma (2003) 'Arrovian Aggregation in Economic Environments: How Much Should We Know about Indifference Surfaces?', *Journal of Economic Theory*, in press.

Fleurbaey, M. and A. Trannoy (2003) 'The Impossibility of a Paretian Egalitarian', *Social Choice and Welfare*, vol. 21, pp. 243–64.

Gale, D. (1967) 'On Optimal Development in a Multi-sector Economy', *Review of Economic Studies*, vol. 34, pp. 1–18.

Kirman, A. P. and D. Sondermann (1972) 'Arrow's Theorem, Many Agents and Invisible Dictators', *Journal of Economic Theory*, vol. 5, pp. 267–77.

Kolm, S. C. (1972) *Justice et équité* (Paris: CNRS).

Koopmans, T. C. (1965) 'On the Concept of Optimal Economic Growth', in The Econometric Approach to Development Planning, *Pontificiae Academiae Scientiarum Scripta Varia*, vol. 28, pp. 225–87.

Lauwers, L. (1995) 'Time-neutrality and Linearity', *Journal of Mathematical Economics*, vol. 24, pp. 347–51.

Lauwers, L. (1997a) 'Continuity and Equity with Infinite Horizons', *Social Choice and Welfare*, vol. 14, pp. 345–56.

Lauwers, L. (1997b) 'Rawlsian Equity and Generalised Utilitarianism with an Infinite Population', *Economic Theory*, vol. 9, pp. 143–50.

Lauwers, L. (1998) 'Intertemporal Objective Functions. Strong Pareto versus Anonymity', *Mathematical Social Sciences*, vol. 35, pp. 37–55.

Lauwers, L. and L. Van Liedekerke (1995) 'Ultraproducts and Aggregation', *Journal of Mathematical Economics*, vol. 24, pp. 217–37.

Maniquet, F. and Y. Sprumont (2004) 'Fair Production and Allocation of an Excludable Nonrival Good', *Econometrica*, vol. 72, pp. 627–40.

Monjardet, B. (1983) 'On the use of Ultrafilters in Social Choice Theory', in P. K. Pattanaik and M. Salles (eds), *Social Choice and Welfare* (Amsterdam: North-Holland).

Pazner, E. and D. Schmeidler (1978) 'Egalitarian Equivalent Allocations: New Concept of Economic Equity', *Quarterly Journal of Economics*, vol. 92, pp. 671–87.

Ramsey, F. P. (1928) 'A Mathematical Theory of Savings', *Economic Journal*, vol. 38, pp. 543–59.

Roemer, J. E and R. Veneziani (2003) 'What We Owe Our Children, They Their Children ...', *Journal of Public Economic Theory*, vol. 6(5), pp. 637–54.

Sakai, T. (2003) 'Intergenerational Preferences and Sensitivity to the Present', *Economics Bulletin*, vol. 4, pp. 1–6.

Sakai, T. (2006) 'Equitable Intergenerational Preferences on Restricted Domains', *Social Choice and Welfare*, vol. 21, pp. 47–54.

Shinotsuka, T. (1998) 'Equity, Continuity, and Myopia: A Generalization of Diamond's Impossibility Theorem', *Social Choice and Welfare*, vol. 15, pp. 21–30.

Suzumura, K. and T. Shinotsuka (2003) 'On the Possibility of Continuous, Paretian and Egalitarian evaluation of Infinite Utility Streams', Mimeo, Institute of Economic Research, Hitotsubashi University.

Svensson, L. G. (1980) 'Equity among Generations', *Econometrica*, vol. 48, pp. 1251–6.

Van Long, N. (2007) 'Toward a Just Savings Principle', ch.15 in this volume.

Varian, H. (1974) 'Efficiency, Equity and Envy', *Journal of Economic Theory*, vol. 9, pp. 63–91.

11

Person-Affecting Paretian Egalitarianism with Variable Population Size*

Bertil Tungodden and Peter Vallentyne

11.1 Introduction

Where there is a fixed population (i.e., one whose existence does not depend on what choice an agent makes), the deontic version of *anonymous Paretian egalitarianism* holds that an option is just if and only if (1) it is anonymously Pareto optimal (i.e., no feasible alternative has a permutation that is Pareto superior), and (2) it is no less equal than any other anonymously Pareto optimal option. We shall develop and discuss a version of this approach for the variable population case (i.e., where who exists does depend on what choice an agent makes). More specifically, we develop and discuss it in the context of a *person-affecting* framework – in which an option is just if and only if it wrongs no one according to certain plausible conditions on wronging.

The general framework

We assume that, for any given option, there is a finite number of possible people who exist in that option. Moreover, we restrict our attention to cases where there is no uncertainty concerning the outcomes of choices.

To fully specify an egalitarian theory, one must specify the type of benefits that it seeks to equalize. Throughout the chapter, however, we leave open the relevant conception of benefit (resources, primary goods, brute luck, well-being, etc.). References to a person being worse off than another should be understood in terms of the relevant benefits.

We assume, for the sake of argument, that benefits are ratio scale measurable and fully interpersonally comparable. The exact informational requirements, however, depend on which version of Paretian egalitarianism (which we here develop) is adopted. For some versions, it is sufficient to have

* For helpful comments, we thank Walter Bossert, Nils Holtug, Wlodek Rabinowicz, Melinda Roberts, the participants at the Hakone Conference, and the participants at the Philosophical Aspects of Social Choice and Welfare Conference, Caen, France (2005).

ordinal measurability and comparability, combined with a norm level stating whether a life is worthwhile living or not. For other versions, a fully comparable ratio scale of benefits is needed. Since we do not focus on any specific version of Paretian egalitarianism, for generality, we assume that benefits are ratio scale measurable with full interpersonal comparability.

We are concerned with the assessment of the justice of alternatives, where alternatives are possible objects of choice (e.g., actions or social policies). Alternatives may have all kinds of features: they generate a certain distribution of benefits, satisfy or violate various rights, involve various intentions, and so on. In what follows, we assume that the only relevant information for the assessment of justice is the benefit distribution that an alternative generates. More formally, we assume:

Benefitism: Alternatives can be identified with (and thus their justice assessed solely on the basis of) their benefit distributions.

Benefitism is a generalization of welfarism. Although it does not assume that welfare (understood narrowly as subjective well-being) is all that matters, it does assume that justice supervenes on individual benefits. If two alternatives generate the same distribution of benefits, then they have the same status with respect to justice. Given Benefitism, we can identify an alternative with the benefit distribution that it generates, and in what follows we do so for simplicity.

We also assume that the set of distributions generated by the set of possible alternatives is *rich* in the following sense:

Domain Richness: For any logically possible benefit distribution X, there is an alternative that generates that distribution.

This condition rules out, for example, the possibility that, where there are just three people, the distribution <3,7,9> (3 to the first person, 7 to the second, 9 to the third) is not one of the alternatives. All logically possible benefit distributions are among the alternatives. This is not to say that all are part of any given *feasible* set (the alternatives that are open to an agent on a given occasion). Of course, there are lots of logically possible benefit distributions that are not feasible on a given occasion. The claim here is about the range of benefit distributions that can be assessed by justice. The condition holds that such judgements can be made for all logically possible distributions. We believe that this is a highly plausible condition. Benefit distributions here play the role of test cases for the theory of justice. All logically possible test cases – assuming, as we shall, a finite population – are admissible.

We also impose the following assumption on the set of feasible sets.

Existence of Individually Best Feasible Option: For any given feasible set, for each individual, there is a maximum feasible benefit.

This rules out feasible sets where one or more person's benefits are unbounded (i.e., can be greater than any standard number) and where everyone's benefits are bounded but one or more person's benefits has no maximum value (e.g., 1/2, 3/4, 7/8, ...). Making sense of rational and moral choice in such cases is very difficult and we shall not attempt to do so here.

Benefitism, Domain Richness and Existence of Individually Best Feasible Option are assumed throughout the chapter, and thus we do not state these conditions explicitly when reporting the results.

Because we are appealing to egalitarian considerations, we need to make explicit some uncontroversial assumptions about the nature of equality. We assume:

Perfect Equality: A distribution X is more equal than a distribution Y if there is perfect equality among the existents in X and not perfect equality among the existents in Y.

Equality Weak Anonymous Contracting Extremes: A distribution X is more equal than a distribution Y, if some permutation of distribution X can be obtained from Y by (1) transferring a fixed amount of benefits from the uniquely best-off person to the uniquely worst-off person – but still leaving each the uniquely best-off and the uniquely worst-off person, respectively, and (2) making no changes in benefits to anyone else.[1]

Equality Acyclicity: If, for distributions $X_1, \dots X_n$, X_1 is more equal than X_2, X_2 is more equal than X_3, ... and X_{n-1} is more equal than X_n, then X_n is not more equal than X_1.

These are each quite uncontroversial. Perfect Equality says, for example, that <2,2,2> is more equal than <1,2,2>. Equality Weak Anonymous Contracting Extremes (which is a weakening of the anonymous version of the well-known Pigou–Dalton condition) says, for example, that <2,5,8> is more equal than <1,5,9>. Equality Acyclicity is a weakened version of transitivity for equality. If X is more equal than Y, and Y is more equal than Z, it allows (unlike transitivity) that Z may be equally good as X or that the two are incomparable.

Because we do not assume that the equality relation is complete, throughout 'a most equal anonymously Pareto optimal option' should be understood as 'is anonymously Pareto optimal and no such option is more equal'. Thus, if there is some incompleteness in the equality relation, an option could still be judged a most equal option, even if it is not at least as equal as all other options.

Justice can be understood in *axiological* terms – what is at least as just as what (i.e., in terms of a justice ranking relation) – or in *deontic* terms – what is just (permitted by justice) relative to a set of feasible alternatives (i.e., in terms of a justice choice function). This latter approach does not attempt to provide a global ranking of alternatives. Instead, it attempts simply to determine which of any given set of feasible alternatives are just. It is well known that if a certain kind of contraction and expansion consistency is required, then the deontic approach is equivalent to the axiological approach. We doubt, however, that contraction consistency is a requirement of justice,[2] and hence that these two approaches are equivalent.

Justice can be understood in different ways, but we here understand it as concerned with what is *owed* to individuals in the sense of what is required to avoid wronging them. We thus assume:

Person-Affecting: An option is just if and only if it wrongs no one.

Person-Affecting would be a controversial thesis if it were a thesis about moral permissibility generally. It would claim that there are no *impersonal wrongs* (wrongs that wrong no one). Although one of us (Vallentyne) is inclined to defend this view, we do not here presuppose it. Instead, we simply limit our attention to justice as what we owe each other (including ourselves). So understood, Person-Affecting is simply a definition of our topic. If there are impersonal wrongs, then any account of justice so understood is an incomplete account of morality. A full account of moral permissibility would then need to deal with the further question of what things are impersonally wrong and how they should be traded-off with personal wrongs.

Nonetheless, person-affecting justice is in itself an important moral topic. A common view is that it is permissible for the state (or private citizens) to forcibly restrict the liberty of citizens only when it is necessary to prevent them from wronging others. Prevention of impersonal wrongs is deemed an insufficient justification for forcibly restricting freedom. Justice in our sense thus provides the basis for assessing the legitimacy of state restrictions of liberty. Of course, if it is not legitimate for the state to restrict a person's liberty to prevent her from wronging *herself* (e.g., suicide), then our account of justice would need to be modified by excluding wrongs to oneself. Such a modification is straightforward once one identifies who the agent is in a given choice situation. For simplicity, however, we leave this modification aside.[3]

We shall also make the following two assumptions, which have been insightfully developed and defended by Roberts (1998, 2002) in the context of a person-affecting framework:

Non-Existence: A person is not wronged by an option if she does not exist under that option.

Best Feasible: A person is not wronged by an option if it is a best feasible option for her.[4]

Non-Existence states that individuals are not wronged by an option if they do not exist under that option. Possible individuals, that is, have no claims to come into existence. It is worth noting here that throughout we understand existence, relative to an option, in an atemporal way. Anyone who existed in the past, exists at the time of choice, or exists in the future, if a given option is adopted, is deemed to exist under that option.

Best Feasible states that a person is not wronged by an option if it is the best feasible option for her. Our assumption of Benefitism (the view that justice is solely concerned with the benefits people get) ensures that whether a person is wronged is determined by the distribution of benefits (as opposed to non-benefits considerations). It leaves open, however, whether a person could be wronged even by the best feasible option for her. Best Feasible rules this out. One might object that in some such cases a person might still be wronged because her best feasible option is still not good enough (e.g., not enough for a decent life, or not enough to give her what she deserves). This objection makes sense if one is concerned with *ideal* justice, that is, with what justice requires ideally, independently of practical constraints of what is possible at the time of choice. We shall, however, limit our attention to *practical* justice which takes feasibility constraints as given, and asks what should be done in that situation. So understood Best Feasible is clearly plausible.[5]

We shall also assume:

No Prohibition Dilemmas: In any choice situation, at least one option is just.

This condition would be controversial if we were concerned with ideal justice, which does not take feasibility constraints into account. We are, however, considering practical justice, which takes such constraints as given, and asks what should be done. Even from this perspective, one could argue that sometimes nothing is just because nothing is good enough. We shall here, however, limit our focus to *comparative* practical justice, according to which justice is purely a matter of comparing favourably in the relevant respects with the feasible alternatives (e.g., being at least as favourable in the relevant respect as all (or 90 per cent) of the feasible alternatives [which is always possible], as opposed to giving everyone an adequate level of benefits [which is not always possible]). Comparative practical justice always satisfies No Prohibition Dilemmas.

Call a framework *basic person-affecting* if it imposes Person-Affecting, Non-Existence, Best Feasible, and No Prohibition Dilemmas. Our task in this chapter is to develop and defend a version of Paretian egalitarianism in the context of a basic person-affecting framework. It is worth noting that Roberts (1998) also invokes a principle that gives priority to benefits to those who exist in both of two alternatives over benefits to those who exist in only one. We address this issue later in the chapter. To start with, however, we do not invoke any such assumption.

11.2 Fixed population

We here introduce a Paretian egalitarian theory that seems promising and is fully consistent with the basic person-affecting framework in the fixed population case (where the same people exist no matter what choice is made). In the next section, we show that this theory is inconsistent with the basic person-affecting framework where there is a variable population (where who exists depends on what option is chosen), and we show how the theory can be revised so as to be fully consistent with the framework without altering its original judgements in the fixed population case.

Before stating the egalitarian theory that we develop, we need to introduce some definitions. An option is *Pareto superior* to another if and only if it makes someone better off and everyone else at least as well off. An option is *Pareto optimal* if and only if no feasible option is Pareto superior. An option is a *permutation* of another option if and only if it has the same distribution of benefits except perhaps with people occupying different positions in the distribution (e.g., <2,1> is a permutation of <1,2>). An option is *anonymously Pareto superior* to another just in case it is Pareto superior to the other or to some permutation of the other. An option is *anonymously Pareto optimal* just in case no feasible option is anonymously Pareto superior to it. Thus, for example, <3,1> is anonymously Pareto superior to <1,2>, and if these are the only two feasible alternatives, then <3,1>, but not <1,2> is anonymously Pareto optimal – even though <2,1> is not feasible. Anonymous Pareto optimality entails Pareto optimality but not vice-versa.

Where there is a fixed population, the following theory seems fairly plausible:

Fixed Population Anonymous Paretian Egalitarianism (FP-APE): An option is just if and only if it is a most equal anonymously Pareto optimal option.

This theory holds that a certain kind of efficiency – anonymous Pareto optimality – is prior to egalitarian considerations. An outcome is just only if it is efficient in this sense. If there are several options that are efficient, then only those that are the most equal among them are just. Of course, the theory is controversial. Many would reject the relevance of equality to justice. Some might accept its relevance, but hold that it is more limited (e.g., limit the role of equality to breaking ties in total benefits). We do not attempt to defend this condition here. Our task is to extend this theory to the variable population case in the context of a basic person-affecting framework. (See Tungodden and Vallentyne, 2005) for some general results on Paretian egalitarianism in the fixed population case.)

The rest of this section records some observations that are rather straightforwardly true in the fixed population case, but which will turn out to fail in the variable population case. To start, it will be instructive to note that FP-APE can be characterized in terms of the following two conditions:

Anonymous Strong Pareto: An option X is unjust if it is not anonymously Pareto optimal.

Weak Egalitarianism Injustice: An option X is unjust if there is a feasible alternative Y that is anonymously Pareto optimal and more equal than X.

Anonymous Strong Pareto strengthens the standard Pareto efficiency requirement by further requiring that even Pareto optimal options be judged unjust if one of their permutations is not Pareto optimal. The strengthening introduces a rather uncontroversial way of solving some of the cases where there is a conflict of interest in the population. Weak Egalitarian Injustice imposes an egalitarian requirement on how to solve the remaining cases of conflicts.

We now note some observations. For brevity, let us say that a theory is *the most permissive theory consistent with a given set of conditions* just in case the theory judges just every option judged just by any other theory that is consistent with the conditions.[6] Consider then:

Observation 1: In the fixed population case, FP-APE is the most permissive theory of justice consistent with the conjunction of Anonymous Strong Pareto and Weak Egalitarianism Injustice.

The proof is straightforward, and hence omitted. We now note that, in the fixed population case, FP-APE is fully consistent with the basic person-affecting framework.

Observation 2: In the fixed population case, given Person-Affecting, FP-APE is consistent with the conjunction of Best Feasible, Non-Existence, and No Prohibition Dilemmas.

Observation 2 can be established as follows. Given that the result only covers a fixed population, it is trivially true that FP-APE is consistent with Non-Existence. Moreover, consider any option X that is the best feasible option for someone that exists. If X is judged just by FP-APE, then Person-Affecting implies that no one is wronged in this alternative, which is consistent with Best Feasible. If X is not judged just by FP-APE, then it is not the most equal anonymously Pareto optimal option. Hence, there is someone who is worse off in this option than in the most equal anonymously Pareto optimal option. FP-APE is consistent with a theory of wronging that states that the person who is worse off in X than in the most equal anonymously Pareto optimal option is wronged in X and the person for which X is the best feasible option is not wronged in X. This theory of wronging is consistent with Best Feasible. No Prohibition Dilemmas is satisfied because (1) there is always at least one anonymously Pareto optimal option, and (2) given that

(a) (as indicated above) 'most equal' is stipulatively understood as 'no option is more equal', and (b) Equality Acyclicity holds, there is always at least one most equal anonymously Pareto optimal option.

In sum, the above results show that, for a fixed population, FP-APE is characterized by Anonymous Strong Pareto and Weak Egalitarian Injustice and is fully consistent with the basic person-affecting framework. As we shall now see, the latter is not the case when we move to the variable population case.[7]

11.3 Variable population

In the variable population case, the people who exist under one option need not be the same as those who exist under another. We shall use '*' to denote non-existence. Thus, in the feasible set {<3,*,2>, <2,4,*>}, the first person exists in both options, the second person exists only in the second option, and the third person exists only in the first option.

There are several issues that need to be clarified if anonymous Paretian egalitarianism is to be applied in the variable population case. First, how do we understand equality? Second, how do we define an anonymously Pareto optimal option?

We impose no controversial assumptions about how to understand equality when the population size is variable. We assume that equality is measured only among those who exist. Thus, for example, we assume that <2,2,*> is perfectly equal, whereas <2,2,0> is not.

With respect to the notion of anonymous Pareto optimality in the variable population case, we first need to make clear how to compare existence with non-existence. We assume that, for a given individual, (1) for any world in which she does not exist, there is some world (not necessarily accessible in a given choice situation) in which she exists that is equally good for her, and (2) any two worlds in which she does not exist are equally good for her. The first assumption is not uncontroversial, but we believe it to be plausible. A world in which an individual receives sufficiently large benefits (e.g., a world that is full of happiness for her) is better for her than any world in which she does not exist, and that any world in which she does not exist is better for her than a world in which she receives sufficiently low negative benefits (e.g., a world full of pain and suffering for her). It is thus plausible to assume that there is some intermediate level of benefits that is equally good for her as non-existence. For a greater defence, see Holtug (2001, 2005). The second assumption is plausible, since the only feature of worlds in which a person does not exist that is relevant for how good that world is for her is her non-existence. Given these two assumptions, we scale benefits so that the zero point is the level of benefits for which it is equally good to exist with those benefits than to not exist at all. Thus, we assume that <2,1> is better for the second person than <2,*>, and that <2,*> is better for her than <2,−1>.

Given this understanding of when an individual is better off, the most natural understanding of Pareto optimality holds that <*,3> is not Pareto optimal when <1,3> is feasible. This is because <1,3> makes the first person better off and the second person no worse off. We shall understand Pareto optimality (and superiority) in this sense. We use, that is, the usual definition of Pareto optimality but combine it with the assumption that non-existence is equally valuable with existence with no benefits.

Next, how is a permutation to be understood for the definition of *anonymous* Pareto optimality? The most natural conception, which we shall adopt, simply treats * (non-existence) as one more value. Thus, <2,3,*> is a permutation of <*,3,2>, but <0,3,2> is not.

With these understandings, we can now show that, in the variable population case, FP-APE is not consistent with the basic person-affecting framework.

Observation 3: In the variable population case, given Person-Affecting, FP-APE does not satisfy the conjunction of Best Feasible and Non-Existence.

To prove the result, consider the feasible set <*,3,1> and <2,2,*>. Both are anonymously Pareto optimal and <2,2,*> is more equal. Hence, FP-APE judges <2,2,*> as just and <*,3,1> as unjust. Given Person-Affecting, this implies that someone is wronged in <*,3,1>. This, however, entails that the conjunction of Non-Existence and Best Feasible is violated. This is because Non-Existence entails that person 1 is not wronged in <*,3,1> and Best Feasible entails that persons 2 and 3 are not wronged in <*,3,1>.

The problem, however, is not merely with FP-APE. We now note:

Observation 4: In the variable population case, Person-Affecting, Best Feasible and Non-Existence are jointly incompatible with each of Anonymous Strong Pareto and Weak Egalitarian Injustice.

The incompatibility with Weak Egalitarian Injustice is illustrated by the example given above. The incompatibility with Anonymous Strong Pareto can be seen by considering the feasible set consisting of <*,1,5>, and <5,1,5>. Best Feasible and Non-Existence entail that no one is wronged in <*,1,5> and Person-Affecting then entails that this option is just, which violates Anonymous Strong Pareto (since <5,1,5> is Pareto superior).

Thus, we need to weaken our Paretian and egalitarian conditions in order to make them compatible with the basic person-affecting framework. Call an option, X, *person-affecting anonymously Pareto optimal* just in case there is no feasible option Y that (1) is anonymously Pareto superior to X and (2) *makes someone existing in X better off.* In the feasible set consisting of <*,1,5> and <5,1,5>, only the second is anonymously Pareto optimal, but both are person-affecting anonymously Pareto optimal (since no anonymously Pareto superior option makes anyone existing in <*,1,5> better off). Consider, then:

Person-Affecting Anonymous Strong Pareto: An option X is unjust if it is not person-affecting anonymously Pareto optimal.

Person-Affecting Weak Egalitarianism Injustice: An option X is unjust if some person-affecting anonymously Pareto optimal option is more equal and makes someone existing in X better off.

In the fixed population case, these two conditions are equivalent to their original counterparts. In the variable population case, however, they are strictly weaker. Neither is violated in our above examples. Person-Affecting Weak Egalitarian Injustice is silent for the feasible set consisting of <*,3,1> and <2,2,*>. Although both are anonymously Pareto optimal – and hence person-affecting anonymously Pareto optimal – and <2,2,*> is more equal than <*,3,1>, the former does not make anyone existing in the latter better off. Likewise, Person-Affecting Anonymous Strong Pareto is silent for the feasible set consisting of <*,1,5>, and <5,1,5>. Although the latter is anonymously Pareto superior to the former, it does not make anyone existing in the former better off.

Consider, then:

Person-Affecting Anonymous Paretian Egalitarianism – Version 1 (PA-APE1): An option, X, is just if and only if (1) X is a person-affecting anonymously Pareto optimal option, and (2) no other such option is more equal and makes someone existing in X better off.

This theory holds, for example, that all three options are just in the feasible set consisting of <*,3,1,*>, <2,*,*,3>, and <2,2,*,*>. Only the second is anonymously Pareto optimal, but all three are person-affecting anonymously optimal (since no other feasible option is both anonymously Pareto superior and makes someone existing in the former better off). Moreover, although the third is more equal than the other two, it does not make anyone existing in the other two better off. Hence, all three are judged just.

We now note:

Observation 5: In the variable population case, given Person-Affecting, PA-APE1 is consistent with the conjunction of Non-Existence, Best Feasible, No Prohibition Dilemmas, Person-Affecting Anonymous Strong Pareto, and Person-Affecting Weak Egalitarian Injustice.

The proof of this observation is as follows:

(1) To see that PA-APE1 is consistent with the conjunction of Non-Existence and Best Feasible, it suffices to note that PA-APE1 is compatible with a theory of wronging that holds than an option X wrongs a person if and only if (a) she *exists in X*, (b) X is *not* a person-affecting anonymously Pareto

optimal option, and (c) some person-affecting anonymously Pareto optimal option is more equal and makes her *better off* than X does. This theory of wronging entails that no person is wronged by an option if she does not exist under that option (Non-Existence) and that a person is not wronged by an option that is the best feasible option for her (Best Feasible).

(2) To see that PA-APE1 satisfies No Prohibition Dilemmas, it suffices to note that (given Equality Acyclicity) there is always at least one option that is person-affecting anonymously Pareto optimal option and for which no other such option is more equal and makes someone existing in X better off.

(3) Finally, PA-APE1's satisfaction of Person-Affecting Anonymous Strong Pareto and Person-Affecting Weak Egalitarian Injustice follows trivially from its definition.

Although PA-APE1 is consistent with the basic person-affecting framework, we believe that it fails to capture some of the spirit of a person-affecting approach. Consider the feasible set consisting of <3,2,*>, <*,2,1>, and <1,3,*>. Only the first is person-affecting anonymously Pareto optimal. The second option is ruled out because the third option is anonymously Pareto superior and makes the second person better off. The third option is ruled out because the first is anonymously Pareto superior and makes the first person better off. Thus, PA-APE1 judges only the first option just. Why, however, should we think that <*,2,1> is unjust? Assuming that <1,3,*> is unjust, everyone existing in <*,2,1> is at least as well off as under every *just* option (since <3,2,*> is the only other possibly just option). More generally, we believe that the following condition is plausible in the context of the person-affecting approach:

No Just Improvements: An option does not wrong an individual if all feasible alternatives that make her better off are unjust.

In the above example, the feasible options are <3,2,*>, <*,2,1>, and <1,3,*>. No Just Improvements says that, if <1,3,*> is judged – by other conditions – unjust, then <*,2,1> wrongs no one. Option <1,3,*> is the only option that makes someone in <*,2,1> better off. Thus, if the former is unjust, it is not possible to make anyone existing in <*,2,1> better off except by choosing an unjust option. No Just Improvements requires that, in this case, no one is wronged by <*,2,1>.

No Just Improvements is similar to Best Feasible. Both say that an individual is not wronged if no 'admissible' option makes her better off. Best Feasible takes all feasible options to be admissible. No Just Improvements, on the other hand, takes options to be admissible only if they are just (on the basis of other conditions). Because it takes a more restrictive view of what is admissible, No Just Improvements entails Best Feasible, but not vice-versa.

In what follows, then, we shall replace Best Feasible by the stronger No Just Improvements.

We can formally note that the above example establishes:

Observation 6: In the variable population case, given Person-Affecting, PA-APE1 violates No Just Improvements.

Indeed, the problem is more general:

Observation 7: In the variable population case, given Person-Affecting, Person-Affecting Anonymous Strong Pareto and Person-Affecting Weak Egalitarian Injustice are each incompatible with No Just Improvements.

The conflict with Person-Affecting Anonymous Strong Pareto is established by the above feasible set consisting of <3,2,*>, <*,2,1>, and <1,3,*>. The conflict with Person-Affecting Weak Egalitarian Injustice can be seen by considering the feasible set consisting of <3,2,*>, <*,1,4>, and <*,5,0>. All three are person-affecting anonymously Pareto optimal. Person-Affecting Weak Egalitarian Injustice judges the third unjust (because, by Equality Weak Anonymous Contracting Extremes, the second is more equal and makes the third person better off) and also judges the second unjust (because the first is more equal and makes the second person better off). Given Person-Affecting, however, this violates No Just Improvements, since everyone who exists in <*,5,0> is at least as well off as under <3,2,*>, which is the only other possibly just alternative.

We believe that No Just Improvements is a plausible condition on justice and we shall therefore assume it in what follows. Call a framework *expanded person-affecting* just in case it satisfies No Just Improvements (and not merely Best Feasible), as well as Person-Affecting, Non-Existence, and No Prohibition Dilemmas. Thus, we must weaken our Pareto and equality conditions even further so as to make them compatible with this expanded framework. Let us say that two options are *anonymously Pareto incomparable* just in case neither is anonymously Pareto superior to the other and neither is a permutation of the other. Consider:

Conditional Person-Affecting Anonymous Strong Pareto: If, for any options X and Y in a given feasible set, (1) option *X is just*, and (2) X is anonymously Pareto superior to Y and makes someone existing in Y better off, then Y is not just.

Conditional Person-Affecting Weak Egalitarian Injustice: If, for any options X and Y in a given feasible set, (1) option *X is just*, and (2) X is anonymously Pareto incomparable to Y, more equal than Y, and makes someone existing in Y better off, then Y is not just.

We show that these two conditions are jointly compatible with the expanded person-affecting framework by appealing to the following theory, which we believe to be eminently plausible. To formulate this theory concisely, we introduce the term *recursively person-affecting most equal Pareto optimal option*, which is defined as follows, where the *unresolved set* is initially the feasible set and then sequentially modified as follows:

(1a) Determine which options are most equal anonymously Pareto optimal options relative to the unresolved set. These options are judged recursively person-affecting most equal Pareto optimal options and are removed from the unresolved set.

(1b) Determine which options have at least one existing person who is worse off than under some option judged to be a recursively person-affecting most equal Pareto optimal option by the previous step. These options are judged *not* to be recursively person-affecting most equal Pareto optimal options and are removed from the unresolved set.

(2) Repeat steps (1a) and (1b) in order until the unresolved set is empty.

(3) An option is a recursively person-affecting most equal Pareto optimal option if and only if so judged by this procedure.

We propose, then:

Person-Affecting Anonymous Paretian Egalitarianism-Version 2 (PA-APE2): An option is just if and only if it is a recursively person-affecting most equal Pareto optimal option.

We shall illustrate the above definition and the resulting theory with reference to the feasible set consisting of <5,7,*>, <9,3,*>, <*,9,3>, <9,*,2>, and <*,8,4>. In the first round, <5,7,*> is judged just because, given Equality Weak Anonymous Contracting Extremes, it is the most equal anonymously Pareto optimal option and <9,3,*> is judged unjust because it makes the second person worse off than <5,7,*>. The unresolved set at this point consists of <*,9,3>, <9,*,2>, and <*,8,4>. In the second round, by Equality Weak Anonymous Contracting Extremes, <*,8,4> is judged just because it is the most equal anonymously Pareto optimal option relative to the unresolved set, and <*,9,3> and <9,*,2> are judged unjust because they each make the third person worse off than under <*,8,4>. Thus, PA-APE2 judges only <5,7,*> and <*,8,4> just. This satisfies No Just Improvements (given Person-Affecting), since each of the other three options makes at least one person worse off than under at least one of these two just options.

We now note that PA-APE2 is consistent with the expanded person-affecting framework:

Observation 8: In the variable population case, given Person-Affecting, PA-APE2 is consistent with the conjunction of Non-Existence, No Just Improvements (and Best Feasible), No Prohibition Dilemmas, Conditional Person-Affecting Strong Pareto, and Conditional Person-Affecting Weak Egalitarian Injustice.

The proof is as follows, where Person-Affecting is assumed throughout:

(1) To see that PA-APE2 is consistent with Non-Existence, it suffices to note that PA-APE2 is compatible with a theory of wronging that holds that an option X wrongs a person if and only if (a) she *exists in* X, (b) X is *not* a recursively person-affecting most equal Pareto optimal option, and (c) some recursively person-affecting anonymously Pareto optimal option is more equal and makes her *better off* than X does.
(2) To see that PA-APE2 is consistent with No Just Improvements, it suffices to note that an option is only judged unjust by PA-APE2 on the basis of a comparison with an option that is judged just.
(3) To see that PA-APE2 satisfies No Prohibition Dilemmas, it suffices to note that (given Equality Acyclicity) there is always at least one most equal person-affecting anonymously Pareto optimal option.
(4) To see that PA-APE2 satisfies Conditional Person-Affecting Strong Pareto consider any feasible set and any two options X and Y thereof, where (a) option *X is just*, and (b) X is anonymously Pareto superior to Y and makes someone existing in Y better off. Since option X is considered just, by step (1a) of the recursive procedure, X must have been the most equal anonymously Pareto optimal options relative to the unresolved set at some step of the procedure. If option Y was part of this unresolved set, then it would be ruled unjust according to step (1b). If Y was not part of this unresolved set, then it would have been judged just or unjust by a previous step in the procedure in relation to a larger unresolved set. However, option X would also have had to be part of this larger unresolved set (since it was judged just only at a later step), and thus option Y could not be an anonymously Pareto optimal option in this larger unresolved set. In sum, option Y has to be judged unjust in relation to one of the two unresolved sets, which is consistent with Conditional Person-Affecting Strong Pareto.
(5) To see that PA-APE2 satisfies Conditional Person-Affecting Weak Egalitarian Injustice, consider any feasible set and any two options X and Y, where (a) option *X is just*, and (b) X is anonymously Pareto incomparable to Y, more equal than Y, and makes someone existing in Y better off. By exactly the same line of reasoning as above, we can show that the recursive procedure has to judge Y as unjust, which is consistent with Conditional Person-Affecting Weak Egalitarian Injustice.[8]

One might wonder whether PA-APE2 is the most permissive theory consistent with the above conditions. The following example shows that this is not so. Consider a theory, PA-APE2*, that makes exactly the same judgements as PA-APE2 for all feasible sets except the one consisting of <3,3,*>, <*,2,4>, <2,*,4>, and <4,2,*>. Here PA-APE2 judges only the first just, whereas PA-APE2*, we stipulate, judges only the last three just. Given Observation 8, and the stipulation that PA-APE2* makes the same judgements as PA-APE2 for all other feasible sets, it follows that PA-APE2* is consistent with all the conditions of the observation for all other feasible sets. Now note that, for the above feasible set, PA-APE2* is consistent with the view that the first option is unjust because it wrongs only the first person. For this set, then, the theory is consistent with No Existence, No Just Improvements, No Prohibition Dilemmas, and the two conditional conditions. So, for this set, PA-APE2* judges just some options that PA-APE2 does not. This establishes that PA-APE2 is not *the* most permissive theory consistent with those conditions.

Although PA-APE2 is not *the* most permissive theory consistent with the conditions of the above observation, it might nonetheless be *a maximally permissive* theory consistent with these conditions, where this means that no other theory that consistent with these conditions (1) judges just every option that it judges just, and (2) also judges just some option that it judges unjust. For illustration of this notion, suppose that theories T1, T2, and T3 satisfy a given set of conditions, and, relative to the feasible set {X,Y,Z}, T1 judges only X just, T2 judges only X and Y just, and T3 judges only Y and Z just. In this case, T1 is not *a maximally permissive* theory consistent with the conditions (since, for this feasible set, the set of options judged just by T2 is a strict superset of those judged just by T1). T2 and T3 may, however, each be *a maximally permissive* theory consistent with those conditions (if they make suitable judgements for other feasible sets). Neither, however, is *the* most permissive theory consistent with the conditions, since each judges some option just that the other judges unjust in the same feasible set.

We now note:

Observation 9: In the variable population case, PA-APE2 is a maximally permissive theory of justice consistent with the conjunction of Conditional Person-Affecting Anonymous Strong Pareto and Conditional Person-Affecting Weak Egalitarianism Injustice.

Above we proved that PA-APE2 is consistent with these two conditions. Here we prove that it is a maximally permissive theory consistent with these conditions. Consider any option, X, judged to be unjust in a given step of the recursive procedure. It is judged unjust in that step if and only if (1) it is not a most equal anonymously Pareto optimal option relative to the unresolved set for that step, and (2) some such option, Y, makes someone in X better off. It follows that Y must be either anonymously Pareto superior to X (which would imply that X is not an anonymously Pareto optimal option) or anonymously Pareto incomparable and more equal (which would imply that it is not a most equal anonymously Pareto optimal option). Given that Y is judged just and makes someone in X better off, in the former case Conditional Person-Affecting Anonymous Strong Pareto requires that X be judged unjust, and in the latter case, Conditional Person-Affecting Weak Egalitarian Injustice requires that X be judged unjust. Thus, given the options judged just by PA-APE2, all theories consistent with Conditional Person-Affecting Anonymous Strong Pareto and Conditional Person-Affecting Weak Egalitarian Injustice must judge all the remaining options unjust. Hence, no other theory consistent with the conditions of the observation (1) judges just every option judged just by PA-APE2, and (2) also judges just some option judged unjust by PA-APE2.

One problem remains that we need to address: Should benefits to individuals who will exist no matter what choice one makes have priority over benefits to those who will exist only if certain choices are made? We now turn to this issue and related issues.

11.4 Gratuitous deprivation

Consider the feasible set consisting of just <9,9,9,9,*> and <1,1,1,1,1>. Is the second option just? According to PA-APE2, both are just, since both are most equal anonymously Pareto optimal options. Many, however, would argue that the second option is not just on the ground that benefits to those who will exist no matter what option is chosen (the first four people in this example) have a certain kind of priority over benefits to those who exist only if certain options are chosen (the fifth person in this example). If the choice is simply between giving those who will exist no matter what very good lives, or creating an extra person with the result that everyone will have a low quality life, it seems unjust to bring the extra person into the world – at least where (1) both are most equal anonymously Pareto options and (2) everyone that exists in the world without the extra person is as well off as is feasible.

The following condition captures a version of this intuition:

Ultra Weak Gratuitous Deprivation: An individual is wronged by an option X if (1) she exists in all feasible options, and (2) there is an option, Y, such that (a) Y is a most equal anonymously Pareto optimal option, (b) Y makes her better off, and (c) everyone who exists in Y is as well off as is feasible.

This condition holds that <1,1,1,1,1> is unjust relative to the feasible set consisting of <9,9,9,9,*> and <1,1,1,1,1>. In this case, both are most equal anonymously Pareto optimal options, but the condition requires that a certain priority be given to the benefits of those who will exist no matter what choice is made. This priority, however, is very weak. First, the condition is silent when the option that is better for the definite existents is not a most equal anonymously Pareto optimal option. For example, it is silent for the feasible set consisting of <1,*,3>, <4,0,*>. Here, although the second option is better for the first person (the only definite existent), it is not a most equal anonymously Pareto optimal. Second, the condition is silent, when even one person is not as well off as possible. For example, it is silent for the feasible set consisting of <1,1,1,1,1>, <9,9,9,9,*>, and <10,0,0,0,0>. Here, although the second option is a most equal anonymously Pareto optimal option and better for all definite existents than <1,1,1,1,1>, the latter is not judged unjust by this condition. This is because the first person is not as well off as possible in <9,9,9,9,*>. The condition only applies when some definitely existing people are better off and makes *everyone* (definitely existing or not) as well off as is feasible.

One of us (Vallentyne) is inclined to accept Ultra Weak Gratuitous Deprivation (indeed something much stronger), but one of us (Tungodden) is inclined to reject it as an added requirement for the expanded person-affecting framework. To see why one might reject it, consider the feasible set consisting of <100,*,*,*,*,*> and <99,99,99,99,99>. Ultra Weak Gratuitous Deprivation holds that the first person is wronged by the second option, and thus, given Person-Affecting, this entails that the second option is unjust. More generally, Ultra Weak Gratuitous Deprivation holds that providing even a very small benefit to just one person who definitely will exist takes priority over large benefits to many more people who exist only if certain choices are made – as long as the former is a most equal anonymously Pareto option. Many people will find that implication difficult to accept.

We do not attempt to resolve this issue. Below we propose a modification to PA-APE2 if some gratuitous deprivation condition is accepted. For the record, however, we briefly note several ways that Ultra Weak Gratuitous Deprivation can be strengthened. Each of these is endorsed by one of us (Vallentyne) and rejected by one of us (Tungodden).

To start consider:

Weak Gratuitous Deprivation: An individual is wronged by an option X if (1) she exists in X, and (2) there is an option, Y, such that (a) Y is a most equal anonymously Pareto optimal option, (b) Y makes her better off, and (c) everyone who exists in Y is as well off as is feasible.

This is like the original condition except that it merely requires that the individual exist *in the given option* rather than that she *exist in all feasible options*. The revised condition thus does not give priority to definite existents as such. Instead, it rules out (roughly) adding people to the world when it would have been possible to add only a proper subset of them and make the members of the subset better off in a certain way. Consider, for example, the feasible set consisting of <1,3,*>, <3,*,1>, and <*,1,1>. The original condition is silent because there are no definite existents. The revised condition, however, judges the first option unjust (because the second option is a most equal anonymously Pareto optimal option, makes the first person, who exists in both, better off, and makes everyone as well off as feasible).

One further strengthening is to drop the requirement in (2a) that the 'dominating' option be a most equal anonymously Pareto optimal option and merely require that everyone existing in both of the options be at least as well off in the 'dominating' option:

Moderate Gratuitous Deprivation: An individual is wronged by an option X if (1) she exists in X, and (2) there is an option, Y, such that (a) everyone who exists in both X and Y is at least as well off in Y as in X, (b) Y makes her better off, and (c) everyone who exists in Y is as well off as is feasible.

Unlike the Weak Gratuitous Deprivation, this judges the first option unjust in the feasible set consisting of <1,*,3>, <4,0,*> – even though <4,0,*> is not a most equal anonymously Pareto optimal option.

A final strengthening is to replace the requirement in (2c) that everyone in Y be as well off as is feasible with the requirement that this be so for those who exist in Y but not in X:

Strong Gratuitous Deprivation: An individual is wronged by an option X if (1) she exists in X, and (2) there is an option, Y, such that (a) everyone who exists in both X and Y is at least as well off in Y as in X, (b) Y makes her better off, and (c) everyone who exists in Y but not in X is as well off as is feasible.[9]

Unlike the above conditions, this judges the first option unjust in the feasible set consisting of <1,*,3>, <4,0,*>, and <5,*,1>. The above conditions are silent about the first option because neither the second nor the third option makes all existents as well off as feasible. The revised condition, however, judges <1,*,3> unjust, because <4,0,*> makes the first person (the only shared existent) better off and makes the second person (the only person who exists in the second but not the first) as well off as possible. The fact that the <4,0,*> does not make the first person (who exists in both) as well off as possible is not deemed relevant.

Unfortunately, we cannot here resolve the issue of whether any of these conditions should be accepted. The important point is that, if we accept at least Ultra Weak Gratuitous Deprivation, then we must modify PA-APE2. As it stands, that theory says that, relative to the feasible set consisting of <9,9,9,9,*> and <1,1,1,1,1>, both options are just (since both are most equal anonymously Pareto optimal options). Ultra Weak Gratuitous Deprivations, on the other hand, requires that <1,1,1,1,1> be judged unjust.

We suppose that, if any gratuitous deprivation condition is imposed, it will be one of the above. For brevity, let us say that a condition C on gratuitous deprivation is *admissible* just in case it is either one of the above conditions or 'the empty condition' that deems that no one is wronged by imposing no requirements. We say that, relative to a given feasible set, an option *C-gratuitously deprives* a person just in case she is wronged according to C.

The most natural revision – which is our final formulation – is the following:

Person-Affecting Anonymous Paretian Egalitarianism – with no C-gratuitous deprivation (PA-APE-C): An option is just if and only if, *relative* *to those feasible options that C-gratuitously deprive no one,* it is a recursively person-affecting most equal Pareto optimal option.

This is just like PA-APE2, except that prior to beginning the recursive procedure, it first eliminates options that C-gratuitously deprive someone.

PA-APE-C is fully consistent with the expanded person-affecting framework combined with No Just Improvements and any gratuitous deprivation condition C:

Observation 10: In the variable population case, given Person-Affecting, PA-APE-C is consistent with the conjunction of any admissible condition C on gratuitous deprivation, Non-Existence, No Just Improvements (and hence Best Feasible), and No Prohibition Dilemmas.

The proof is as follows, where Person-Affecting is assumed throughout:

(1) It follows straightforwardly that PA-APE-C is consistent with any condition C on gratuitous deprivation.
(2) It follows from Observation 8 that PA-APE-C is consistent with the conjunction of Non-Existence, No Just Improvements (and hence Best Feasible) and No Prohibition Dilemmas relative to the set of options that gratuitously deprive no one. We now address whether this is so relative to the entire feasible set.
(3) To see that PA-APE-C is consistent with No Prohibition Dilemmas for the entire feasible set, we have to show first that for any admissible C, there will always be a non-empty set of alternatives that C-gratuitously deprives no one. Consider any admissible version of C and suppose that someone is C-gratuitously deprived in X. This means that there is an option Y where (i) this person exists in X, and (ii) there is an option, Y, such that (a) everyone who exists in both X and Y is at least as well off in Y as in X, (b) Y makes her better off, and (c) everyone who exists in Y but not in X is as well off as is feasible. If Y gratuitously deprives no one, then we have established that at least one option gratuitously deprives no one. If Y does gratuitously deprive someone, then it must gratuitously deprive someone who exists in both X and Y (since all other individuals are as well off as feasible). Hence, there must exist some W where (a) everyone who exists in both Y and W is at least as well off in W as in Y, (b) W makes some person (who exists in both X and Y) better off than in Y, and (c) everyone who exists in W but not in Y is as well off as is feasible. Given that we have assumed Existence of Individually Best Feasible Option, it follows that there exists such an alternative W that does not gratuitously deprive anyone. Hence, given Equality Acyclicity, it follows that there exists at least one most equal person-affecting anonymously Pareto optimal option – thereby satisfying No Prohibition Dilemmas.

(4) It is obvious that PA-APE-C is also consistent with Non-Existence for the entire feasible set, given that any admissible condition C on gratuitous deprivation only states that someone is wronged in an option if they exist in that option.
(5) To see that No Just Improvements is satisfied for the entire feasible set, we have to establish that for any alternative X that gratuitously deprives someone, there exists a just alternative that makes someone existing in X better off. If, for any admissible condition C on gratuitous deprivation, X C-gratuitously deprives someone, then, given (3), there is an alternative W that does not gratuitously deprive anyone and such that (i) everyone who exists in both X and W is better off in W and (ii) everyone who only exists in W is as well off as is feasible. Given PA-APE-C, if W is not just, then there exists some Z that is just and makes someone existing in W better off in Z than in W. However, those who are better off in Z than in W must be among those who exist in X, since the others in W are as well off as is feasible. This implies that there is someone existing in X who is better off in Z (since everyone existing in X is better off in W and some of them are better off in Z) and hence that No Just Improvements is satisfied.

We now note:

Observation 11: In the variable population case, PA-APE-C is a maximally permissive theory of justice consistent with the conjunction of an admissible condition C on gratuitous deprivation, Conditional Person-Affecting Anonymous Strong Pareto, and Conditional Person-Affecting Weak Egalitarianism Injustice.

The proof is straightforward: Any theory satisfying an admissible condition C on gratuitous deprivation has to judge unjust – as does PA-APE-C – options that C-gratuitously deprive someone. Consider now the set of all remaining options. With respect to this set PA-APE-C makes the same judgements as PA-APE2. Observation 9 establishes that PA-APE2 is a maximally permissive theory consistent with Conditional Person-Affecting Anonymous Strong Pareto and Conditional Person-Affecting Weak Egalitarianism Injustice. It follows that PA-APE-C is a most permissive theory consistent with the three conditions.

We believe that PA-APE-C is the most plausible way of adapting anonymously Paretian egalitarianism to the expanded person-affecting framework. Obviously, PA-APE-C requires more scrutiny before it can be accepted with any confidence. Our task here, however, is simply to formulate and motivate a promising person-affecting version of anonymous Paretian egalitarianism. We believe that PA-APE-C is such a theory. Before concluding, we note how the recursive person-affecting approach used by PA-APE-C can be generalized to other theories (such as utilitarianism).

11.5 A generalization: recursively person-affecting theories

A generalized form of the recursively person-affecting procedure that we invoked to define PA-APE2 can be used to make any theory consistent with the expanded person-affecting framework – as expanded to include No Just Improvements, and perhaps an admissible condition on gratuitous deprivation. Moreover, we suggest that this way of making a theory consistent with the expanded person-affecting framework is the most plausible way of doing so.

Consider any theory of justice, T, and any admissible condition, C, on gratuitous deprivation. Define *recursively person-affecting* T-C as follows, where the *unresolved set* is initially the feasible set and then sequentially modified as follows:

(1) Judge unjust all options that C-gratuitously deprive someone and remove them from the unresolved set.
(2a) Determine which options are judged *just* by T relative to the unresolved set. These options are judged just and are removed from the unresolved set.
(2b) Determine which options have at least one existing person who is worse off than under some option judged to be recursively person-affecting T-just by the previous step. These options are judged *not* to be just and are removed from the unresolved set.
(3) Repeat steps (2a) and (2b) in order until the unresolved set is empty.
(4) Recursively person-affecting T judges an option just if and only if it is so judged by this procedure.

Consider, for example, (total) utilitarianism, saying that the minimal set of just alternatives consist of all alternatives with the greatest total utility. It is incompatible with the expanded person-affecting framework. We suggest that the most plausible modification of utilitarianism consistent with the expanded person-affecting framework is recursively person-affecting utilitarianism. For illustration, consider the feasible set consisting of <*,2,5>, <3,3,*>, <3,*,2>, and <*,4,*>. Utilitarianism judges only the first just and, given Person-Affecting, this violates the conjunction of Non-Existence and No Just Improvements (which requires that <3,3,*>, and <*,4,*> each also be judged just, since, in each, each existing person is at least as well off as under the only just option, <*,2,5>). Recursively Person-Affecting Utilitarianism, however, satisfies both these conditions in this case. Assuming that no condition on gratuitous deprivation is imposed, then, in the first round, no judgements are made. In the second round, <*,2,5> is judged just and <3,*,2,> is judged unjust. In the third round, <3,3,*> is judged just and nothing is judged unjust. In the fourth and final round, <*,4,*> is judged just. This is consistent with the expanded person-affecting framework.

Suppose now that Strong Gratuitous Deprivation is imposed. In this case, in the first round, <*,2,5> and <3,3,*> are judged unjust (because each gratuitously deprives the second person in comparison with <*,4,*>). In the second round, <3,*,2> is judged just and nothing is judged unjust. In the third round, <*,4,*> is judged just and nothing is judged unjust. Thus, only <3,*,2> and <*,4,*> are judged just, and inspection shows that this is consistent with the expanded person-affecting framework.

We now note the following general result:

Observation 12: For any theory of justice, T, that satisfies No Prohibition Dilemmas, and any admissible condition, C, on gratuitous deprivation, recursively person-affecting T-C is consistent with the conjunction of Person-Affecting, Non-Existence, No Just Improvements (and Best Feasible), No Prohibition Dilemmas, and condition C.

The proof follows straightforwardly from the proofs of Observation 8 and Observation 10.

The generalized recursive person-affecting procedure thus converts any theory of justice into one that satisfies the expanded person-affecting framework. We believe, moreover, that it does so in a particularly plausible manner. Thus, for example, if one is committed to utilitarianism in the fixed population case, and endorses the expanded person-affecting framework with admissible condition C on gratuitous deprivation, then, we suggest, one should endorse recursively person-affecting utilitarianism-C. We do not, however, attempt to argue for this claim here.

11.6 Conclusion

We have assumed the person-affecting framework, which is defined by Person-Affecting, Non-Existence, and Best Feasible. We further suggested that Best Feasible should be strengthened to No Just Improvements. Finally, we assumed that, in the fixed population case, FP-APE is correct. In the variable population case, however, FP-APE is ruled out by the person-affecting framework. More generally, Strong Pareto and Weak Egalitarianism Injustice are each ruled out. We suggested that each should be weakened in a certain way and showed that PA-APE2 is consistent with the conjunction of all these conditions.

We also discussed the issue of gratuitous deprivation, but came to no conclusion on this difficult issue. We suggested, however, that, if some admissible condition, C, of gratuitous deprivation is imposed, then PA-APE2 should simply be applied to the set of options that satisfy that condition. More exactly, we suggested that the following view is plausible (PA-APE-C): An option is just if and only if, *relative to those feasible options that C-gratuitously deprive no one*, it is a recursively person-affecting most equal Pareto optimal option.

Finally, we suggested that, for any theory of justice, T, the most plausible way of modifying it to make it compatible with the expanded person-affecting framework – augmented by No Just Improvements, and an admissible gratuitous deprivation condition, C – is to apply a generalized version of the recursive procedure invoked by PA-APE-C to obtain recursively person-affecting T-C.

We close with a few comments on how the person-affecting framework – and thus any recursively person-affecting theory – deals with various versions of the (deontic) repugnant conclusion. Because the repugnant conclusion raises particular problems for utilitarianism (and similar aggregative theories), we shall focus on recursively person-affecting utilitarianism for illustration.

Suppose that one has the choice between (1) an option where many people have good lives, and (2) an option where those people do not exist, many more other people exist with lives just barely worth living, and the total benefits are greater. Suppose, for example, that the choice is between <9,9,*,* . . . [20 times] . . .* ,*> and <*,* , 1, 1 . . . [20 times], . . . 1, 1 >. (For simplicity, we use small numbers of people for illustration, but the idea can be made more striking by supposing that each number represents a billion people.) The first has two people with a total of 18 and the second has 20 different people with a total of 20. Because all individuals existing in the first option are as well off as feasible, the person-affecting framework ensures that it is judged just. Thus, a strong form of the repugnant conclusion is avoided. Justice does not *require* one to choose the option producing a highly populated, but fairly bleak world. Even recursively person-affecting utilitarianism agrees with this judgement: It judges both just.

The person-affecting framework, however, is subject to a weak version of the repugnant conclusion in cases such as the above. The framework – that is, the conjunction of Person-Affecting, Best Feasible (or with the stronger No Just Improvements), and No-Existence – requires that, in the above case, justice *allow* one to choose the option producing highly populated, but fairly bleak world. This is because, in this particular kind of case, everyone in that world is as well off as is feasible. The judgement that it is just to choose such an option (even if it is also just not to do so) will strike many as bizarre. Within the person-affecting framework, however, it is inevitable and natural. Who is wronged by such a choice? Not the individuals who exist with the bleak, but worth living, lives. Their lives are better than non-existence, which is the only alternative. Nor are individuals who do not exist wronged. Hence, no one is wronged and the option is indeed just. Of course, it might be impersonally wrong to choose such an option, but we have set aside that issue in this chapter.

Let us now consider a repugnant conclusion case in which some people exist under more than one option. Suppose, for example, that the choice is between <9,9,*,*…[20 times] …*,*> and <1,1,…[22 times], …1,1>. As above, the first option has two people with a total of 18. Because both are still as well off as feasible, the person-affecting framework rightly requires that the first option be judged just. This time, however, the second option has 22 people with a total of 22 *and two of these people also exist under the first option*. It is thus no longer true that everyone in the second option is as well off as feasible, and the person-affecting framework no longer requires that the second option be judged just. Nonetheless, both PA-APE2 and recursively person-affecting utilitarianism judge the second option just. Again, this avoids the strong version of the repugnant conclusion (since the second option is not required by justice), but it faces the weak version thereof (since the second option is permitted by justice). If, however, we further add at least Ultra Weak Gratuitous Deprivation (which requires that the second option be judged unjust), then even the weak form of the repugnant conclusion is avoided in these kinds of cases. Although we have left open whether Ultra Weak Gratuitous Deprivation should be endorsed, it is clear that it provides an important way of avoiding certain versions of the repugnant conclusion.

In sum, the person-affecting framework, we believe, has the resources to avoid the main problematic versions of the repugnant conclusion. Because we find some version of anonymous Paretian egalitarianism attractive, we have focused on it. We believe that PA-APE3-C is the most plausible version thereof that is compatible with the expanded person-affecting framework. Our more general claim, however, is that the generic recursively person-affecting procedure is a plausible way of converting any theory into one that is consistent with the expanded person-affecting framework. Obviously, many of the judgements invoked in the chapter are controversial. We hope nonetheless that we have at least established that the person-affecting framework should be taken seriously and that there are promising ways of developing anonymous Paretian egalitarianism – and other theories – within this framework.

Notes

1. This is an anonymous version of a condition introduced and discussed by Vallentyne (2000a).
2. More specifically, we deny that Alpha is required: If an alternative is judged just relative to a given feasible set, then it is also judged just from any subset containing it. For criticism of this condition, see Tungodden and Vallentyne (2005) and Sen (1993).
3. The person-affecting idea can also be expressed in terms of axiological justice: A distribution is less just than another only if it is worse for someone. We are, however, skeptical that the axiological person-affecting approach is promising in the variable population case. Any such approach, we believe, will have to be radically incomplete so as to avoid generating cycles of betterness (i.e., where X_1 is better than X_2, which is better than $X_3, \ldots$ which is better than X_n, which is better than X_1,).

4. As stated, Best Feasible overlaps with Non-Existence when non-existence is a best feasible option for an individual (e.g., both say that person two is not wronged by <2,*> when the only alternative is <2,−3>). To avoid this overlap, we could have restricted Best Feasible to only cover options where a person exists, but we have not done so since this would require cumbersome expressions below.
5. Note that Best Feasible is compatible with holding that a person is wronged by being created with a life not worth living, when non-existence is feasible.
6. We here assume that theories of justice are identical if, for all possible feasible sets, they judge the same options just. Thus, there is only one theory that is a maximally permissive theory consistent with given conditions.
7. The literature on variable population ethics is extensive. See, for example, Arrhenius (2005), Blackorby, Bossert and Donaldson (2005), Broome (2004), Holtug (2005), Roberts (1998, 2002), and Vallentyne (2000b).
8. For the record, we note that PA-APE2 violates the weak anonymity condition that requires that, if an option and a permutation thereof are each feasible, then either both are just or neither is. To see that violation, consider the feasible set consisting of <3,3,*>, <3,1,*>, and <*,3,1>. PA-APE2 judges only the first and third option just. The violation of anonymity is effectively unavoidable within the person-affecting framework.
9. This principle is tentatively endorsed by Roberts (1998) in note 48 of ch. 2. It is a strengthening of her official principle D*, which is the same except that clause (4) says that no one exists in Y but not in X.

References

Arrhenius, Gustaf (2005) 'The Paradoxes of Future Generations and Normative Theory', in J. Ryberg and T. Tännsjö (eds), *The Repugnant Conclusion* (Dordrecht: Springer), pp. 201–18.

Blackorby, Charles, Walter Bossert, and David Donaldson (2005) *Population Issues in Social-Choice Theory, Welfare Economics and Ethics* (Cambridge, MA: Cambridge University Press).

Broome, John (2004) *Weighing Lives* (Oxford: Oxford University Press).

Holtug, Nils (2001) 'On the Value of Coming into Existence', *Journal of Ethics*, vol. 5, pp. 361–84.

Holtug, Nils (2005) 'Person-affecting Moralities', in Jesper Ryberg and Torbjorn Tannsjo (eds), *The Repugnant Conclusion* (Dordrecht: Springer), pp. 129–61.

Roberts, Melinda (1998) *Child versus Childmaker: Future Persons and Present Duties in Ethics and the Law* (New York: Rowman and Littlefield).

Roberts, Melinda (2002) 'A New Way of Doing the Best That We Can: Person-based Consequentialism and the Equality Problem', *Ethics*, vol. 112, pp. 315–50.

Sen, Amartya (1993) 'Internal Consistency of Choice', *Econometrica*, vol. 61, pp. 495–521. Reprinted in Amartya Sen, *Rationality and Freedom* (Cambridge, MA: Harvard University Press, 2002).

Tungodden, Bertil and Peter Vallentyne (2005) 'On the Possibility of Paretian Egalitarianism', *Journal of Philosophy*, vol. 102, pp. 126–54.

Vallentyne, Peter (2000a) 'Equality, Efficiency, and Priority for the Worse Off', *Economics and Philosophy*, vol. 16, pp. 1–19.

Vallentyne, Peter (2000b) 'Critical Notice of *Child versus Childmaker: Future Persons and Present Duties in Ethics and the Law* by Melinda Roberts', *Nous*, vol. 34, pp. 634–47.

Part IV
Sustainability and Human Development

12

Intergenerational Justice and Sustainability Under the Leximin Ethic*

John E. Roemer

12.1 Introduction

Let me begin by proposing that we think about intergenerational justice from the viewpoint of equality of opportunity. According to the equal-opportunity view, a person should be compensated if his welfare is low due to circumstances beyond his control, but not if it is low due to actions or choices that we (our society) thinks he should be held responsible for. The word I use to denote the second set of actions/choices is 'effort'. Consider, now, a standard economic growth model with a representative agent at each generation. The society consists of the set of all agents who will ever live, one representing each generation. Clearly, the pre-eminent circumstance, for an individual in this society, is the date at which he is born. So if we apply the equal-opportunity view, and if we assume that individuals are identical except for their birthdates, we would have to say that justice requires an intertemporal resource allocation which enables all individuals, regardless of their date of birth, to acquire the same level of welfare. We must, however, be interested in efficiency as well as equity, so the just and efficient allocation of resources is that one which enables all individuals, regardless of their birth date, to achieve the same level of welfare, where that level is the highest possible such level. Even this, however, may not be Pareto efficient – it may be possible to render all individuals (weakly) better off than at that any equal-welfare distribution. And so we finally say that the just and efficient allocation of resources is that one which renders the worst-off representative agent (across generations) as well off as possible: the intergenerational maximin distribution.

* I thank Roberto Veneziani for our discussions and his careful comments. My interest in studying the possibility of increasing welfare along the leximin path has been kindled by discussion and work with Joaquim Silvestre. I am grateful to Koichi Suga for correcting errors in a previous draft, and to conference participants for comments.

I believe this is probably what intergenerational equity requires, although the short argument I have just given can surely be challenged. Note, in particular, that I do not discount the welfare of future generations. There might well be grounds for such discounting if there is a probability that future generations may not exist, but my feeling is that this justification for discounting is very much overplayed by economists. We do not have much doubt that the next ten or even fifty generations will exist, and so, at least, we should not discount *their* welfare. But fifty is almost infinity.

The kind of well-being of an individual in this society with which I am concerned is a function only of her own consumption, not of the consumption of her descendants. One may think of this assumption in one of two ways: either individuals indeed *do not* worry about their descendants, or, more persuasively, the kind of well-being with which we, the impartial ethical observer, should be concerned, is the *standard of living* of individuals, rather than the kind of happiness they get when contemplating their children's and grandchildren's lives, which is properly called *welfare*. *We*, the ethical observer, will take care of the children and grandchildren.

We now assume that there is a natural environment, to be thought of as a resource that can both be used in production to produce consumption goods, or can be enjoyed in its pristine state. Think of nature as a forest, which can either be harvested for timber to build houses, or used for hiking and recreation. The forest has a natural rate of growth, or regeneration, and let us assume that it is depleted only when used for timber, but not for hiking. Clearly, there is a rate at which wood can be harvested, so that the forest would remain of constant size. Call this rate the *sustainable* rate of harvesting the forest; that rate is not of any particular ethical significance, because we have not yet related it in any way to human welfare, if our ethics concern only the welfare of humans.

Sustainable has at least two meanings in the environmental literature. One meaning is anthropocentric: a path of resource use is sustainable if, over time, human welfare does not decrease. The other meaning is greener, and is the one I used just above: a path of resource use is sustainable if the stock of the resource does not decrease, or does not decrease to zero. I will be concerned with both meanings in this chapter.

A more modern image than one of 'forest' and 'housing' would be to imagine the natural resource as the biosphere, which can either be enjoyed as a health- and life-giving resource, or can be depleted to produce manufactured consumption goods. I will retain the names 'forest' and 'houses' for the sake of convenience.

We assume that the standard of living (or, for this chapter, her welfare) of the individual at a given date is a function of two arguments, the consumption goods (houses) that she produces from the part of the forest she harvests as timber at that date, and of the hiking she does in the pristine forest that remains uncut. Houses depreciate fully at each generation, and so the only bequest one generation leaves to the next are the remaining forest, and its technological knowledge.

Suppose that members of our society strive to improve their lives, and to this end, they engage in research which produces technological improvements – that is, at least under normal conditions, there will be an exogenous rate of productivity increase in the use of timber to make houses. Suppose that each generation costlessly passes down its technology to the next one. Thus, technological striving produces a positive externality for future generations.

We are interested in the allocation of use of the natural resource for this society. We may denote an allocation as a function $R:\mathbb{Z}_+ \to \mathbb{R}_+$, which specifies the size of the forest at date $t \in \mathbb{Z}_+$ that remains in its uncut state, and is used for hiking. Clearly, specifying a function R is equivalent to specifying the amount that will be harvested at each date, and used for the production of houses. The question is: What is the function R that maximizes the minimum level of welfare across all generations?

Formally, we can state this problem as follows. At date t, if I is the harvest of the forest as timber, then housing in the amount $\alpha^t I$ can be produced. The utility function of each agent is $u(R, H)$ where R is the recreational use of the forest and H is housing. We assume that the recreational value of the forest is just measured by the size of the forest left uncut. The natural rate of growth of the forest is ρ. Then we wish to solve:

$$\begin{aligned} &\max_{I(\cdot),R(\cdot)} \min_{t} u(R(t), \alpha^t I(t)) \\ &s.t. \quad R(t) = \rho R(t-1) - I(t), \quad t \geq 1 \qquad \text{Programme (1)} \\ &R(t) \geq 0 \text{ for all } t \geq 1 \end{aligned}$$

The data of the problem are the sequence $\{\alpha^t | t = 1, \ldots\}$, the natural growth factor of the forest ρ, and $R(0)$, the initial size of the forest. I will assume that $u(0, H) = u(R, 0) = u^{\min}$ for any R, H, and $u'(0, H) = u'(R, 0) = \infty$, so that the optimal solution to Programme (1) must entail positive consumption of both housing and hiking at every date. Note that no resources are consumed in technological innovation.

12.2 The leximin solution

I must first be somewhat more precise. There may, indeed, be many maximin solutions. What we really are interested in is the lexicographic minimum solution, which I define as follows.

Consider the sequence of programmes $\mathbf{P}(n)$, $n = 1, 2, \ldots$, defined as follows:

$$\mathbf{P}(n) \quad \begin{aligned} &\max_{I(\cdot),R(\cdot)} \min_{t \geq n} u(R(t), \alpha^t I(t)) \\ &s.t. \quad R(t) = \rho R(t-1) - I(t), \quad t \geq n \\ &R(t) \geq 0 \text{ for all } t \geq n. \end{aligned}$$

A path $R^*(t)$ is a *leximin solution* if it solves $\mathbf{P}(n)$ for all $n \geq 1$. If R^* solves $\mathbf{P}(1)$ it is a maximin solution. If it solves $\mathbf{P}(2)$, it is, among all maximin solutions, one which is also maximin for generations two and beyond; and so on.

Consider an intertemporal path R. We will say that generation t is *unconstrained on the path R* if, given $R(t-1)$ and α^t, the value of $R(t)$ is that value which is individually optimal for generation t. In other words, we say an individual in this society is unconstrained if the prescribed path at date t is just the value of the forest that agent t would decide upon were he to optimize selfishly, given his endowment, and ignoring future generations.

If an agent is not unconstrained on a path, then we say he is constrained. We immediately have:

Proposition 1. *In a leximin solution, if an agent is constrained, then he consumes less housing than is individually optimal for him, given his endowment.*

The proof is immediate. If he consumed more housing than were individually optimal, he could decrease his housing consumption to the individually optimal amount, and at the same time increase the forest endowment for future generations, which would render all future generations better off, and the solution could not be leximin.

Now denote by $u^t(R)$ the utility of generation t on a path R. Below, I will often simply write this as u^t.

We have the following:

Proposition 2. *If R is a leximin solution then:*

1. *for all* $t \geq 1$, $u^t(R) \leq u^{t+1}(R)$.
2. *For any t, if* $u^t(R) < u^{t+1}(R)$, *then agent t is unconstrained.*

Proof:

Claim 1. Suppose $u^t > u^{t+1}$, some t. Then the value of $\mathbf{P}(t)$ is equal to the value of $\mathbf{P}(t+1)$. Now have the agent at t consume a little less timber. This raises the endowment of programme $\mathbf{P}(t+1)$, and hence raises its value; it consequently raises the value of $\mathbf{P}(t)$. Hence the original solution was not leximin.

Claim 2. Suppose some agent were constrained. Then by Prop. 1, he could increase his utility by increasing I(t) a small amount. This raises the value of $\mathbf{P}(t)$, and hence the original solution was not leximin. □

Proposition 3. *Suppose that u is strictly quasi-concave. Then there is a unique leximin solution.*

Proof:

1. Let R^* be a leximin solution. By Prop. 2, the minimum utility is achieved at $t=1$. This utility is the same in all leximin solutions. The budget constraint for each generation, in $(R(t),I(t))$ space, has slope -1. There are two possibilities at $t=1$: either the agent's indifference curve is tangential to the budget line $R(1)+I(1)=\rho R_0$ or it cuts the budget line in (at most) two points. If the latter, then we know $I(t)$ is uniquely determined by Prop. 1. If the former, then it is tangential at exactly one point, by strict quasi-concavity. Hence $R(1)$ is uniquely determined.
2. Hence the endowment of $\mathbf{P}(2)$ is uniquely determined for all leximin solutions, and it follows that $R(2)$ is uniquely determined, as above.
3. By induction, *QED*.

From Proposition 2, we see that leximin solutions can be grouped into three classes:

Class 1. $u^t(R) < u^{t+1}(R)$ for all $t \geq 1$
Class 2. $u^t(R) = u^{t+1}(R)$ for all $t \geq 1$;
Class 3. $u^t(R) \leq u^{t+1}(R)$, with some equalities and some inequalities.

I will say that solutions of Class 1 comprise Nirvana: for by Proposition 2, these are *intergenerationally just paths where each generation may optimize selfishly*. In other words, in a state of Nirvana, the needs of future generations place no constraint on earlier generations' use of the forest.

12.3 Intertemporal optimization with leximin utility: a general theorem

In the case when the intergenerational optimization problem is well-behaved (a concave problem), we have a characterization theorem for the leximin path with constant utilities at every date.

Denote the partial derivatives of a function $u(x,y)$, by u_1 and u_2. Define $u^t(R,I) = u(R,\alpha^t I)$.

Theorem 1 *Let* $\{u^t\}$ *be a sequence of increasing, concave utility functions defined on* $\mathbb{R}^2$ *and let* $\{I(t), t=1,2,\ldots\}$, $\{R(t), t=1,2,\ldots\}$ *be non-negative sequences* *satisfying*

$$\textit{for all } t = 1, 2, \ldots \quad R(t) = \rho R(t-1) - I(t).$$

$$\textit{Define } a^t = \frac{u_1^t(R(t), I(t))}{u_2^t(R(t), I(t))} \textit{ and suppose}$$

$$\textit{for all } t = 1, 2, \ldots . \quad a^t = a \leq 1, \textit{ for some constant } a, \textit{ and}$$
$$u^t(R(t), I(t)) = k, \textit{ for some constant } k.$$

Then $\{I(t), R(t)\}$ *solves the programme*:

$$\begin{aligned}
&\max u^1(\rho R(0) - I(1), I(1)) \\
&s.t. \\
&R(t) = \rho R(t-1) - I(t), \quad t = 1, 2, \ldots \\
&u^t(R(t), I(t)) \geq k, t = 2, 3, \ldots
\end{aligned} \tag{P}$$

Proof:

1. A *feasible point* for programme (P) is a pair of non-negative sequences $\{I(t), t=1,2,\ldots\}$, $\{R(t), t=1,2,\ldots\}$ satisfying the constraints of programme (P), for a given, fixed value of k. Observe that the set of feasible points is convex. It follows that (P) is a concave programme.
2. Suppose the claim were false, and that there is a feasible point which gives a higher value than k to the objective. Denote the investment sequence in this dominating solution by $\{I(t) + g^t, t = 1, 2, \ldots .\}$. Let $\{\lambda^t, t = 1, 2, \ldots\}$ be an arbitrary sequence of non-negative numbers, and define the function:

$$J(\varepsilon) = u^1(\rho R(0) - (I(1) + \varepsilon g^1), I^1 + \varepsilon g^1) + \sum_{t=2}^{\infty} \lambda^{t-1}(u^t(R^t(\varepsilon), I(t) + \varepsilon g^t) - k)$$

where $\{R^t(\varepsilon)\}$ is defined recursively by:

$$\begin{aligned}
R^t(\varepsilon) &= \rho R^{t-1}(\varepsilon) - (I(t) + \varepsilon g^t) \\
R^0(\varepsilon) &= R(0)
\end{aligned}$$

J takes values on the extended real line.

Note that J is a concave function and that $J(0) = k$, by hypothesis. According to the claim, $J(1) > k$, since $J(1)$ is the value of programme (P) at the proposed solution $\{I(t) + g^t\}$ plus the sum of non-negative terms. Therefore, if we can demonstrate that there exists a choice of the non-negative sequence $\{\lambda^t\}$ such that the associated function J is maximized at

$\varepsilon = 0$, then in particular, $k = J(0) \geq J(1)$, a contradiction to the claim that $\{I(t)\}$ is not the solution of (P).

3. By computing $R^t(\varepsilon)$ for a few terms we observe that its derivatives are:

$$\frac{\mathrm{d}R^2(\varepsilon)}{d\varepsilon} = -\rho g^1 - g^2$$

$$\frac{\mathrm{d}R^3(\varepsilon)}{d\varepsilon} = -\rho^2 g^1 - \rho g^2 - g^3, \text{ etc.}$$

and so we may compute the derivative:

$$J'(0) = -u_1^1 g^1 + u_2^1 g^1 + \lambda^1 (u_1^2(-\rho g^1 - g^2) + u_2^2 g^2) + \lambda^2 (u_1^3(-\rho^2 g^1 - \rho g^2 - g^3) + u_2^3 g^3) + \ldots \quad (12.3.1)$$

The function u_j^t in this expression is of course evaluated at $(R(t), I(t))$. Gathering terms, we can write the coefficient of g^1 in this expression as:

$$C^1 = u_2^1 - \sum_{t=0}^{\infty} \lambda^t \rho^t u_1^{t+1}, \quad \text{where } \lambda^0 = 1.$$

Let C^t be the coefficient of g^t in expression (12.3.1). If we can find a non-negative sequence $\{\lambda^t\}$ such that $C^t = 0$ for all t, then we will have demonstrated that, for that choice, $J'(0) = 0$. We may write the vanishing of all these coefficients as the following system:

$$C^1 \qquad u_2^1 - u_1^1 = \sum_{t=1}^{\infty} \lambda^t \rho^t u_1^{t+1}$$

$$C^2 \qquad \lambda^1(u_2^2 - u_1^2) = \sum_{t=2}^{\infty} \lambda^t \rho^{t-1} u_1^{t+1}$$

$$C^3 \qquad \lambda^2(u_2^3 - u_1^3) = \sum_{t=3}^{\infty} \lambda^t \rho^{t-2} u_1^{t+1}, \text{ etc.}$$

Let us define $z = u_2^1 - u_1^1$. Then we may write these equations, beginning with the second one, as:

$$C^2 \qquad \lambda^1(u_2^2 - u_1^2) = \frac{z - \lambda^1 \rho u_1^2}{\rho}$$

$$C^3 \qquad \lambda^2(u_2^3 - u_1^3) = \frac{z - \lambda^1 \rho u_1^2 - \lambda^2 \rho^2 u_1^3}{\rho^2}, \text{ etc.}$$

We may solve these equations sequentially for the sequence $\{\lambda^t\}$, giving:

$$\lambda^1 = \frac{z}{\rho u_2^2}, \quad \lambda^2 = \frac{z}{\rho^2 u_2^3}(1-a), \quad \lambda^3 = \frac{z}{\rho^3 u_2^4}(1-a)^2, \text{ etc.}$$

Note that by hypothesis, since $a \leq 1$, it follows that $z \geq 0$, and so all the λ^t are non-negative.

The only condition left to check is the equation associated with C^1 above. Substituting the derived values of the λ^t into that condition, we deduce that what must be verified is:

$$z = za + za(1-a) + za(1-a)^2 + \ldots. \tag{12.3.2}$$

If $z = 0$, this is surely true. If $z > 0$, then dividing (12.3.2) by z immediately shows that (12.3.2) reduces to an identity. Thus, we have produced a sequence of non-negative multipliers for which $J'(0) = 0$. □

12.4 Exogenous technical change: special cases of utility function

I now specialize on particular cases, in order to be able to compute leximin solutions with comparative ease. We assume:

A1. The rate of technological progress is exogenous and constant: for every

$$t \geq 1, \ \frac{\alpha^{t+1}}{\alpha^t} = \gamma, \text{ some } \gamma > 1.$$

The first case I study is:

A2. Cobb-Douglas preferences: $u(R, H) = R^b H^{1-b}, b \in (0, 1)$.
We have:

Theorem 2 *Let **A1** and **A2** hold. Then*:

A. *If* $\rho < \frac{1}{b\gamma^{1-b}}$, *then the leximin solution is given recursively by*

$$R^*(t) = \frac{\rho}{1+\delta} R^*(t-1), \ I^*(t) = \frac{\rho\delta}{1+\delta} R^*(t-1), \quad \text{where } \delta = \rho\gamma^{1-b} - 1.$$

Utilities are equal for all generations, and every generation is constrained.

B. *If* $\rho \geq \frac{1}{b\gamma^{1-b}}$, *then the leximin solution is given recursively by:*

$$R^*(t) = b\rho R*(t-1), \quad I^*(t) = (1-b)\rho R^*(t-1).$$

No generation is constrained, and utility increases at each date.

C. *In case* (A), *the size of the forest decreases monotonically to zero. In case* (B), *the size of forest decreases to zero if* $b < \frac{1}{\rho}$, *and increases without bound if* $b > \frac{1}{\rho}$.

The theorem tells us that sustainability in the green sense of a non-shrinking forest size occurs only if love of the forest (b) is sufficiently great. In particular, in case (B) of the theorem we have Nirvana, but nevertheless the forest may shrink to zero.

Lemma 1 Let $R^*(t)$ be a path at which each agent is unconstrained and $u^t \leq u^{t+1}$ for all t. Then R^* is the leximin solution.

Proof:

u^1 cannot possibly be greater, and so R^* solves $\mathbf{P}(1)$. Hence $R^*(1)$ is determined. By induction, the path is the leximin path. □

Proof of Theorem 2:

1. We prove part B first. We begin by asking: Is there a value γ at which, if all agents optimize selfishly, the path generated will enjoy equal utilities at all dates? If so, by Lemma 1, this is the leximin path for that process of technical change.
2. If t optimizes selfishly then

$$R(t) = b\rho R(t-1)$$

$$I(t) = (1-b)\rho R(t-1)$$

and so

$$I(t) = \frac{1-b}{b} R(t).$$

It follows that

$$\frac{u^{t+1}}{u^t} = \frac{(b\rho R(t))^b (\alpha^{t+1})^{1-b} ((1-b)\rho R(t))^{1-b}}{R(t)^b (\alpha^t)^{1-b} \left(\frac{1-b}{b} R(t)\right)^{1-b}}$$

$$= b\rho\gamma^{1-b} = 1 \Leftrightarrow \rho = \frac{1}{b\gamma^{1-b}}$$

Hence, if the last equation holds, the leximin solution is Nirvana, and utilities are equal at all dates.
3. It follows that if $\rho > \frac{1}{b\gamma^{1-b}}$, then individual optimization at each date generates increasing utilities, which, by Lemma 1, must be the leximin solution.
4. Now suppose $\rho < \frac{1}{b\gamma^{1-b}}$.

Compute that the solution defined in statement A produces equal utilities at all dates, that is:

$$u^t(R(t), I(t)) = \left(\frac{\alpha^1 \delta}{\gamma}\right)^{1-b} \quad \text{for all } t \geq 1,$$

and note that $\frac{\delta}{1+\delta} < 1 - b$, which shows that $I^*(t)$ is less than the individually optimal harvest for agent t.

5. We next verify that the other premise of Theorem 1 holds. Compute that

$$a^t = \frac{u_1^t(R(t), I(t))}{u_2^t(R(t), I(t))} = \frac{b\delta}{1+\delta}.$$

What remains to be verified, then, is only that $b\delta \leq 1 + \delta$. This is equivalent to the statement $b\rho\gamma^{1-b} \leq 1$, which is true by the premise of Part A.

Hence, by Theorem 1, $(I^*(t), R^*(t))$ is a maximin solution of our programme. It now follows, by repeated application of theorem 1, that it is also the leximin solution. □

Suppose that our own natural forest regenerates at a rate of 3 per cent per annum. If a generation is 25 years, then $\rho = 1.03^{25} = 1.85$, and $\frac{1}{\rho} = 0.54$. Is $b < 0.54$? If so, intergenerational justice and green sustainability are incompatible with this utility function.

My second exercise involves a different utility function. One might well say: human beings require some minimal amount of forest. The idea that welfare might remain constant by continually substituting housing consumption for recreation in the forest (health) is ridiculous. So let us replace **A2** by:

A3. $u(R, H) = (R - x_0)^b H^{1-b}$, some $0 < x_0 < \rho R(0)$.

With **A3**, we have a Stone–Geary utility function, with a minimal necessary consumption of forest. The upper bound on x_0 simply assures that it is possible for the first generation to satisfy its minimal need of hiking with the forest that it inherits.

We now have:

Theorem 3 Assume **A3**.

A. *If* $b\rho + (1-b)x_0 < 1$ *then there is a unique process* $\{\hat{\alpha}^t\}$ *of technical progress at which the leximin solution yields constant utilities and unconstrained agents at all dates. Under this process, the size of the forest decreases over time and approaches the value* $\frac{x_0(1-b)}{1-b\rho}$ *in the limit.*

B. *If* $b\rho + (1-b)x_0 < 1$ *and technical progress is more rapid than in* $\{\hat{\alpha}^t\}$ *at all* t*, then all agents are unconstrained, utilities increase over time, and the limit size of the forest is as in* A.

C. *If* $b\rho + (1-b)x_0 > 1$ *then under any process of technical progress the leximin solution has agents unconstrained at every date, and utilities increase over time. If* $b\rho < 1$ *then the size of forest converges to the limit in* A; *if* $b\rho > 1$*, the size of the forest grows without bound.*

Proof:

1. We attempt to compute a value of γ at which selfish optimization at each date produces equal utilities at all dates. Selfish optimization now entails:

$$R(t) = b(\rho R(t-1) - x_0) + x_0 \tag{12.1}$$

$$I(t) = (1-b)(\rho R(t-1) - x_0) \tag{12.2}$$

Note that $I(t) = \frac{1-b}{b}(R(t) - x_0)$.
Hence

$$\frac{u^{t+1}}{u^t} = \frac{(b(\rho R(t) - x_0))^b(\alpha^{t+1})^{1-b}((1-b)(\rho R(t) - x_0))^{1-b}}{(R(t)-x_0)^b(\alpha^t)^{1-b}(\frac{1-b}{b}(R(t)-x_0))^{1-b}} = \tag{12.3}$$

$$\frac{b(\rho R(t) - x_0)\gamma^{1-b}}{R(t) - x_0} = 1$$

$$\Leftrightarrow$$

$$R(t)\{1 - b\rho\gamma^{1-b}\} = x_0(1 - b\gamma^{1-b})$$

For the last equation to hold there are two possibilities: either $R(t)$ is constant over time at the required value, or $1 = b\rho\gamma^{1-b}$ and $x_0 = 0$. In the first case, utilities would increase with time, because of technical progress, which contradicts our hypothesis; and in the second case, we are back in the Cobb-Douglas world. So, if $x_0 > 0$, there is no constant rate of technical progress that will engender the solution we seek.

2. Returning to equation (12.3), and relaxing the assumption of a constant rate of technical progress, we require:

$$(\gamma^{t+1})^{1-b} = \frac{R(t) - x_0}{b(\rho R(t) - x_0)}. \tag{12.4}$$

A path where every agent optimizes and (12.4) holds will have equal utilities. By Lemma 1, this will be the leximin path.

3. We study the process $\{\alpha^t\}$ that would produce (12.4) when each generation optimizes selfishly. From (12.1) and (12.4) we have:

$$(\gamma^{t+1})^{1-b} = \frac{R(t) - x_0}{R(t+1) - x_0}. \qquad (12.4)'$$

   Consequently, a process of technical progress (with $\alpha^{t+1} > \alpha^t$) necessarily involves $R(t+1) < R(t)$.
4. By the recursion (12.1), compute that

$$R(t) = (b\rho)^t + x_0(1-b)(1 + b\rho + \cdots + (b\rho)^{t-1}) \qquad (12.5)$$

   From (12.5), compute that

$$R(t+1) - R(t) = (b\rho)^{t+1} - (b\rho)^t + (b\rho)^t x_0(1-b) < 0$$
$$\Leftrightarrow b\rho + (1-b)x_0 < 1. \qquad (12.6)$$

   So if (12.6) holds, we can define uniquely a process of technical progress for which (12.4′), and hence (12.4), holds. Because *a fortiori* $b\rho < 1$ in this case, we compute from (12.5) that

$$\lim R(t) = \frac{x_0(1-b)}{1-b\rho}. \qquad (12.7)$$

   This proves statement A.
5. An aside: the formula (12.5) only describes $R(t)$ if at each date $\rho R(t) > x_0$. Failing this inequality, the agent would have to consume less than x_0 in hiking. Since the sequence $R(t)$ described in step 4 decreases monotonically to the value in (12.7), it suffices to verify the inequality $\frac{\rho x_0(1-b)}{1-b\rho} \geq x_0$, which is true.
6. Statement B follows immediately. If every generation optimizes selfishly, then (12.1) holds, but now $(\gamma^{t+1})^{1-b} > \frac{R(t)-x_0}{R(t+1)-x_0}$ for all t, and so utilities increase at every date. This is the leximin solution by Lemma 1.
7. Statement C. If agent 1 optimizes then $R(1) = b(\rho R(0) - x_0) + x_0 = b\rho + (1-b)x_0 > 1$.

Hence agent 2 has a greater forest endowment and a better technology than agent 1, and so if he optimizes selfishly, he will be better off than agent 1. We know that, if every agent optimizes then $R(t+1) > R(t)$ (step no. 4), and so an increasing sequence of utilities is generated. By Lemma 1, this is the leximin solution, proving the first part of statement C. Since the size of the forest at date t is given by (12.5), we see that it approaches $\frac{x_0(1-b)}{1-b\rho}$ if $b\rho < 1$ and becomes infinitely large if $b\rho > 1$. □

The hopeful part of Theorem 3 is part C. It says that if x_0 is fairly close to $\rho R(0)$, then the leximin solution entails Nirvana for *any* process of technical progress. This is a bit surprising: one might have thought that if x_0 is 'large', we would be in a situation of scarcity, which would have bad implications for the forest. However, the size of the forest will increase to a finite limit, unless b is large.

The Cobb-Douglas utility function has an elasticity of substitution of unity. Finally, I study the CES utility function where there is more complementarity between hiking and housing than in the Cobb-Douglas function. We might hope that, with these preferences, citizens will not be so willing to deplete the forest. Here is a theorem:

Theorem 4 Let $u(R,H) = (bR^r + (1-b)H^r)^{1/r}$, $r < 0$. Let $\{\alpha^t\}$ be any process of technical progress, and define $p^t = \frac{1-b}{b}(\alpha^t)^r$ and $\varphi(t) = \frac{(p^t)^{\frac{1}{r-1}}}{1+(p^t)^{\frac{1}{r-1}}}$.

A. If the law of motion of technical change is given by

$$\varphi(t+1) = \rho^{\frac{r}{1-r}}\varphi(t)^{\frac{1}{1-r}} \qquad (12.8)$$

then we have Nirvana with constant utilities. We have

$$\lim \varphi(t) = \rho^{-1}, \lim \alpha^t = \alpha^* \equiv \left(\frac{b}{1-b}\right)^{\frac{1}{r}} (\rho - 1)^{\frac{1-r}{r}} \quad \text{and } \lim \frac{R(t+1)}{R(t)} = 1.$$

B. Let $\{\alpha^t\}$ be a process of technological progress such that for large t, $\alpha^t > \alpha^*$. Then $\varphi(t) > \frac{1}{\rho}$ and the forest increases in size at every date. Henceforth, all agents optimize selfishly, and utilities increase.

Proof:

1. If a generation optimizes selfishly, then compute that:

$$R(t) = \varphi(t)\rho R(t-1), \quad I(t) = (1-\varphi(t))\rho R(t-1).$$

2. Hence, if each generation optimizes selfishly then:

$$\left(\frac{u^{t+1}}{u^t}\right)^r = \frac{b(\varphi(t+1)\rho R(t))^r + (1-b)(\alpha^{t+1}(1-\varphi(t+1))\rho R(t))^r}{bR(t)^r + (1-b)(\alpha^t \frac{1-\varphi(t)}{\varphi(t)} R(t))^r}$$

$$= \frac{\rho^r[\varphi(t+1)^r + p^{t+1}(1-\varphi(t+1))^r]}{1 + p^t(\frac{1-\varphi(t)}{\varphi(t)})^r}.$$

By inverting the defining equation of φ, compute that $p^t = (\frac{1-\varphi(t)}{\varphi(t)})^{1-r}$, and substituting this into the last expression, we compute that $\frac{u^{t+1}}{u^t} = 1 \Leftrightarrow \frac{\varphi(t+1)^{1-r}}{\varphi(t)} = \rho^r$ which is equation (12.8) in statement A.

3. Note that $\varphi(t) \in (0, 1)$, and so a limit of $\varphi(t)$ exists. It follows from the recursive definition of φ that $\lim \varphi(t) = \varphi^* = 1/\rho$. The expression in statement A for $\lim \alpha^t$ follows immediately from the definition of φ and p.
4. $\frac{R(t+1)}{R(t)} = \rho\varphi(t) \to \rho\varphi^* = 1$.
5. If $\alpha^t > \alpha^*$ then we have:

$$
\begin{aligned}
&(\alpha^t)^r < \frac{b}{1-b}(\rho - 1)^{1-r} \Rightarrow \\
&p^t < (\rho - 1)^{1-r} \Rightarrow \\
&(p^t)^{\frac{1}{r-1}} > \frac{1}{\rho - 1} \Rightarrow \varphi(t) > \frac{1}{\rho} \Rightarrow \\
&R(t) = \varphi(t)\rho R(t-1) > R(t-1).
\end{aligned}
$$

Thus, selfish optimization yields an increasing size of forest. Hence we have achieved Nirvana with increasing utilities. □

This is our most hopeful result. With CES utility functions, in which $r < 0$, there is a finite value of the technological coefficient α, which, if exceeded, will permit all generations to optimize selfishly along the just path, well-being will increase at every date, and the forest increases in size.

Thus, our preliminary investigation suggests that the most hopeful scenario for consistency between the green view and justice *à la* equality-of-opportunity is that hiking and consumption of produced commodities are more complementary than they are in Cobb-Douglas preferences.

For instance, if $r = -0.5$ and $b = 0.5$ and $\rho = 1.85$, then $\alpha^* = 1.628$.

12.5 Endogenous technical change

In this section, we introduce endogenous technical change.

We would like to study the programme

$$
\begin{aligned}
&\max \quad k \\
&s.t. \\
&u(R(t), H(t)) \geq k, \quad t = 1, 2, \ldots \\
&R(t) = \rho R(t-1) - I(t), \quad t = 1, 2, \ldots \\
&H(t) = (1 - \lambda(t))^a \alpha(t) I(t)^{1-a} \\
&\alpha(t) = (1 + \hat{\gamma}\lambda(t))\alpha(t-1)
\end{aligned}
$$

where $\alpha(0), R(0), \gamma, \rho$ are given. ρ is the natural growth factor of the forest. Think of $\lambda(t)$ as the fraction of the labour force assigned to the R&D industry at date t, the rest working in the housing industry. Production in the housing industry is Cobb-Douglas: in the R&D industry, the single input is labour time, and production is linear in labour.[1] Thus, the rate of technical change, if all labour were used in R&D, is $\hat{\gamma}$. Please note that $\hat{\gamma}$ of this section corresponds to $\gamma - 1$ of previous sections.

The intergenerational leximin solution solves this programme. Moreover, we will show that the solution to this programme must be the leximin solution.

The endogeneity of technical progress makes this a non-concave programme. (Just check that if (λ, R, I) and $(\hat{\lambda}, \hat{R}, \hat{I})$ are two feasible paths, their convex combinations are not generally feasible.) This is a general feature of endogenizing technical change in growth models. Uzawa (1965) solves an intertemporal problem with endogenous technical change (though not in the sustainability framework and not in the leximin framework) by demonstrating that there is a *unique* path which satisfies the necessary conditions for a maximum. We will not, however, attempt to solve this non-concave problem here. Most growth theorists who study endogenous technical change do not solve the general non-concave problem: they restrict themselves to a subset of the set of all feasible paths which is convex (e.g., constant growth paths). This, for instance, is the strategy of Lucas (1988).

There is an important externality here: if generation t invests in R&D, it reaps the benefits from that investment, and the new technological knowledge is passed on free to the future. This feature is embodied in the constraint defining $\alpha(t)$. This externality gives us the possibility of supporting *increasing utilities* on the leximin path – which we study below.

It is worth nothing that it is not immediately apparent how one might capture this externality if one modelled the problem in continuous time. In the continuous case, the technology constraint becomes $\dot{\alpha}(t) = \hat{\gamma}\lambda(t)$. No amount of investment in R&D at time t can increase the value of α at t: all it does is increase the rate at which α increases for people in the future. Hence, generation t will have no selfish motive to invest in R&D in the continuous model. Thus, as we are interested in this kind of externality, it appears that we must use the discrete-time model.

I do not here study the general non-concave problem stated above. I study a simpler programme where the R&D industry is constrained to employ a constant *fixed* fraction of the labour force: $\lambda(t) = \lambda$ for all t.

We solve this programme, and derive k as a function of λ. We then may choose $\lambda \in [0, 1]$ to maximize k. Of course, this does not solve the general non-concave programme, which may well involve varying the fraction of the labour force employed in R&D.

Normalize by setting $R(0) = 1$. Our programme is:

$$\begin{aligned} &\max \quad k \\ &s.t. \\ &\left.\begin{aligned} &u(R(t), (1-\lambda)^a \alpha(t) I(t)^{1-a}) \geq k \\ &R(t) = \rho R(t-1) - I(t) \\ &\alpha(t) = (1 + \hat{\gamma}\lambda)^t \alpha_0 \end{aligned}\right\} P(\lambda) \end{aligned}$$

A *point* is a feasible path {*I(t),R(t)*}. Note this is a concave programme for λ fixed.

Consequently, by the variational method used in the proof of Theorem 1, we can compute shadow prices. (Define the appropriate function $J(\cdot)$ as in the proof of that theorem.) We here search for the characterization of a solution that entails constant utility across time.

If, at a point {*I(t),R(t)*} at which $u^t = k$ for all t, there are non-negative sequences $\{x^t\}$, $\{y^t\}$ such that:

$$\begin{aligned} &1 = \sum_1^\infty x^t \\ &0 = x^t u_1^t - y^t + \rho y^{t+1}, \quad t = 1, 2, \ldots \\ &0 = u_2^t (1-\lambda)^a \alpha(t)(1-a) I(t)^{-a} x^t - y^t, \ t = 1, 2, \ldots \end{aligned}$$

then the point is a solution of the programme.

Now we specialize to the case $u(R,H) = R^c H^{1-c}$, so $u_1^t = c(H/R)^{1-c}$, $u_2^t = (1-c)(R/H)^c$. But $R^c H^{1-c} = k$, and so $u_1^t = \frac{ck}{R(t)}$, $u_2^t = \frac{(1-c)k}{H(t)}$. Thus the stated conditions can be written:

(1) $1 = \sum x^t$

(2) $0 = \dfrac{ckx^t}{R(t)} - y^t + \rho y^{t+1}$

(3) $0 = \dfrac{x^t(1-c)k(1-a)}{I(t)} - y^t$

$\Rightarrow x^t = \dfrac{I(t) y^t}{(1-c)k(1-a)}$

Substituting into (2) we have

(2′) $\rho y^{t+1} = y^t \left(1 - \dfrac{I(t)}{R(t)} \dfrac{c}{(1-c)(1-a)}\right).$

We now try for a solution where $R(t) = A^t$, some $A > 0$. Repeated use of the 'budget' constraint gives

$$I(t) = A^{t-1}(\rho - A), t = 1, 2, \ldots$$

Therefore $\frac{I(t)}{R(t)} = \frac{\rho - A}{A}$ and so (2′) becomes

$$\rho y^{t+1} = y^t \left(1 - \frac{\rho - A}{A} \frac{c}{(1-c)(1-a)}\right).$$

Consequently we must check that:

$$1 \geq \frac{\rho - A}{A} \frac{c}{(1-c)(1-a)} \tag{T1}$$

to guarantee the non-negativity of the y^t.

Denote $Q = \frac{1}{\rho}(1 - \frac{\rho - A}{A}\frac{c}{(1-c)(1-a)})$. Then $y^t = Q^{t-1}y^1$ for $t = 1, 2, \ldots$ since $\frac{y^{t+1}}{y^t} = Q$. Therefore $x^t = \frac{Q^{t-1}y^1 A^{t-1}(\rho - A)}{(1-c)k(1-a)}$ for all t, from (3). So (1) requires that

$$1 = \frac{y^1(\rho - A)}{(1-c)k(1-a)} \sum_{t=1}^{\infty} (QA)^{t-1} \text{ or}$$

$$1 = \frac{y^1(\rho - A)}{(1-c)k(1-a)} \frac{1}{1 - QA}. \tag{12.4}$$

For this series to sum as stipulated, we must have

$$QA < 1. \tag{T2}$$

But we have

$$\begin{aligned} k &= R(t)^c[\alpha(t)I(t)^{1-a}(1-\lambda)^a]^{1-c} \\ &= [A^c(1+\hat{\gamma}\lambda)^{1-c}A^{(1-c)(1-a)}]^t \alpha_0^{1-c} \left(\frac{\rho - A}{A}\right)^{(1-a)(1-c)} (1-\lambda)^{a(1-c)} \end{aligned}$$

which implies that the term in square brackets is unity:

$$A^{c+(1-c)(1-a)}(1+\hat{\gamma}\lambda)^{1-c} = 1 \text{ and so}$$

$$A = (1+\hat{\gamma}\lambda)^{\theta} \text{ where } \theta = \frac{c-1}{c+(1-c)(1-a)}.$$

Note that $\theta < 0$, and so we have $A < 1 < \rho$.

This gives us the formula:

$$k = \alpha_0^{1-c} \left(\frac{\rho - A}{A} \right)^{(1-a)(1-c)} (1 - \lambda)^{a(1-c)}, \tag{12.5}$$

which defines the function $k(\lambda)$.

From (4), we now define:

$$y^1 = \frac{(1 - QA)(1 - c)k(1 - a)}{\rho - A}.$$

Because $A < 1 < \rho$ we have $\rho - A > 0$. Consequently we have solved for non-negative sequences $\{x^t, y^t\}$ that satisfy conditions (1)–(3), subject to verifying (T1) and (T2).

We check (T1) and (T2). (T1) reduces to the statement:

$$\rho < \frac{c-1}{c\theta}(1 + \hat{\gamma}\lambda)^{\theta}.$$

On the other hand, (T2) reduces to the statement

$$A\left(1 + \frac{c}{(1-c)(1-a)}\right) < \rho\left(1 + \frac{c}{(1-c)(1-a)}\right),$$

which is true. Therefore we have the following:

Proposition 4 *Let* $u(R, H) = R^c H^{1-c}$. *Let* $\rho < \frac{c-1}{c\theta}(1 + \hat{\gamma}\lambda)^{\theta}$. *Then the solution to* $P(\lambda)$ *entails* $u^t = k$ *for all* t, *where*

$$R(t) = A^t$$

$$I(t) = A^{t-1}(\rho - A)$$

$$\text{and } A = (1 + \hat{\gamma}\lambda)^{\theta}, \quad \text{where } \theta = -\frac{1-c}{c + (1-c)(1-a)}.$$

k is given by eqn. (12.5).

I began with a feasible point at which $u^t = k$ for all t. The variational method shows that this point is a solution to the programme $P(\lambda)$, under the premise of the proposition, which is a maximin programme. One can see that this must be the leximin solution. This entails looking at the maximin programme which begins at date 2 with the endowment $\rho A = \rho R(1)$. The same argument shows that the solution of that programme entails $u^2 = k$. By induction on the date, we have that the maximin solution at each date, along this path, taking at each date the endowment from the previous date along the path, yields constant utilities. But this means the solution is the leximin solution.

The solution of Proposition 4 is not Nirvana: at every date, the agent would like to consume more of the forest as housing, but he is obliged not to, for the sake of future generations. We next study the conditions for the Nirvana solution to hold, where each generation maximizes selfishly on the leximin path. If selfish choice generates a non-decreasing sequence of utilities, then it is the leximin solution.

If generation 1 maximizes selfishly, then it chooses R to

$$\max R^c[(1-\lambda)^a(1+\hat{\gamma}\lambda)\alpha_0(\rho-R)^{1-a}]^{1-c}$$

which is equivalent to solving

$$\max R^{\frac{c}{c+(1-c)(1-a)}}(\rho-R)^{\frac{(1-c)(1-a)}{c+(1-c)(1-a)}}$$

whose solution is

$$R(1)=\frac{c}{c+(1-c)(1-a)}\rho, I(1)=\frac{(1-c)(1-a)}{c+(1-c)(1-a)}\rho.$$

This gives a value of utility at the first date of:

$$u^1=\beta^c(1-\beta)^{(1-c)(1-a)}\rho^{c+(1-c)(1-a)}(1-\lambda)^{a(1-c)}[(1+\hat{\gamma}\lambda)\alpha_0]^{1-c} \qquad (12.6)$$

where $\beta=\frac{c}{c+(1-c)(1-a)}$. Now the available forest at date 2 is $\beta\rho^2$, and it follows by the same reasoning that if the date 2 agent maximizes selfishly, she chooses:

$$R(2)=\beta^2\rho^2, I(2)=(1-\beta)\beta\rho^2.$$

By substituting these values into her utility function, we can compute that on this path:

$$\frac{u^2}{u^1}=\beta^{c+(1-c)(1-a)}\rho^{c+(1-c)(1-a)}(1+\hat{\gamma}\lambda)^{1-c}. \qquad (12.7)$$

Indeed, this is the ratio for the utilities at any two consecutive dates along the selfish path. We require, then, that this ratio be at least unity, which is equivalent to the statement:

$$\beta\rho\geq(1+\hat{\gamma}\lambda)^{\theta},$$

which in turn says that:

$$\rho\geq\frac{c-1}{c\theta}(1+\hat{\gamma}\lambda)^{\theta}.$$

In summary:

Theorem 5 (Cobb-Douglas utility) *Let* $\lambda\in[0,1]$. *If* $\rho<\frac{c-1}{c\theta}(1+\hat{\gamma}\lambda)^{\theta}$, *then the leximin solution entails constant utilities at each date, with the path given in Proposition 4*. If $\rho\geq\frac{c-1}{c\theta}(1+\hat{\gamma}\lambda)^{\theta}$, *then the leximin path entails selfish maximization by* *the agent at each date, and if this inequality is strict, then utilities increase at each date.*[2]

We now consider the issue of choosing among the leximin paths associated with the various values of λ. Define the function $\rho(\lambda) = \frac{c-1}{c\theta}(1+\hat{\gamma}\lambda)^{\theta}$. Note that this function is decreasing. Define the number λ^* as the solution of the equation

$$\rho = \rho(\lambda^*), \text{ that is, } \lambda^* = \frac{\left(\frac{c\theta\rho}{c-1}\right)^{1/\theta} - 1}{\hat{\gamma}}.$$

Then the theorem tells us that:

- If $\lambda^* \leq 0$, then for every $0 \leq \lambda < 1$, the leximin path is Nirvana;
- if $\lambda^* \geq 1$, then for every $0 \leq \lambda < 1$, the leximin path has constant utilities;
- if $0 < \lambda^* < 1$, then for $0 \leq \lambda \leq \lambda^*$, the leximin path has constant utilities, and for $\lambda^* < \lambda < 1$ the leximin path is a Nirvana path.

We now find the leximin path among all Nirvana paths, should they exist. This is the path among these that maximizes utility at the first date. This utility is given in eqn. (12.6); hence, we must choose λ to maximize $(1-\lambda)^a(1+\hat{\gamma}\lambda)$. The solution of the FOC of this function (after taking logarithms to render it concave) is $\lambda = \frac{\hat{\gamma}-a}{\hat{\gamma}(1+a)}$. We therefore have:

Among Nirvana paths, the leximin solution is given by

$$\lambda_1^{lex} = \begin{cases} \lambda^*, \text{ if } \dfrac{\hat{\gamma}-a}{\hat{\gamma}(1+a)} \leq \lambda^* \\ \dfrac{\hat{\gamma}-a}{\hat{\gamma}(1+a)}, \text{ otherwise.} \end{cases}$$

Among constant-utility paths, the leximin path maximizes $k(\lambda)$, which is to say that it maximizes

$$(\rho(1+\hat{\gamma}\lambda)^{-\theta} - 1)^{1-a}(1-\lambda)^a.$$

The derivative of the logarithm of this concave function is

$$-\left(\frac{a}{1-\lambda} + \frac{(1-a)\theta\rho\hat{\gamma}(1+\hat{\gamma}\lambda)^{-(1+\theta)}}{\rho(1+\hat{\gamma}\lambda)^{-\theta} - 1}\right).$$

If this expression is somewhere zero in the interval $[0, \lambda^*]$, then that value of λ generates the leximin path among all constant-utility paths. If this expression is everywhere positive in this interval, then the leximin path among constant-utility paths occurs at $\lambda = \lambda^*$; if it is everywhere negative in the interval, then the leximin path occurs at $\lambda = 0$.

Finally, the *overall* leximin path, among all paths with constant λ, is found by comparing the two leximin paths just computed.[3] The one which gives a higher utility at date 1 is the overall leximin path. (If they give the same utility at date 1, then the Nirvana path, with increasing utilities, is the leximin path.)

By looking at the formula for λ^*, we see that condition $\lambda^* \leq 0$ is equivalent to the condition

$$1 \leq c(1-\rho\theta) \quad \text{or } \rho \geq 1 + \frac{(1-c)(1-a)}{c}. \tag{12.8}$$

If this is true, then at every $\lambda < 1$ the leximin solution is Nirvana, so the overall leximin solution is Nirvana. It is interesting to note that (12.8) does not involve $\hat{\gamma}$.

We note an interesting fact concerning the fate of the forest along the leximin path. Note that, in the constant-utility paths, $R(t) = A^t$, and since $A < 1$, the forest size approaches zero. However, in the Nirvana paths, we have $R(t) = (\beta\rho)^t$ and so the forest size increases without bound, stays the same size, or decreases to zero, as $\beta\rho$ is greater than, equal to, or less than one, respectively. More precisely, we have:

Theorem 6 (Cobb-Douglas utility) *Fix a value* $0 < \bar{\lambda} < 1$.

A. *The forest sizes increases without bound on the leximin path at* $\bar{\lambda}$ *if and only if it increases without bound on the* λ-leximin *path* for every $\lambda \in (0, 1)$.
B. *If the forest sizes increases without bound on the* $\bar{\lambda}$*-leximin path then utilities increase without bound on the* λ*-leximin path at every* λ.

Proof:

1. We have noted, just above, that the condition for the forest's increasing without bound on the $\bar{\lambda}$-leximin path is that the path be Nirvana *and* that $\beta\rho > 1$. This is equivalent to $\bar{\lambda} > \lambda^*$ and $\frac{c\theta\rho}{c-1} > 1$. But the last inequality implies that $\lambda^* < 0$, and so the inequality $\bar{\lambda} > \lambda^*$ is redundant. Therefore, if the forest increases without bound at $\bar{\lambda}$, then the λ-leximin path is Nirvana for every λ (since $\lambda^* < 0$) and hence the size of the forest increases without bound on every λ-leximin path (because the inequality $\frac{c\theta\rho}{c-1} > 1$ is independent of λ). The converse of statement A is trivial.
2. The premise of statement B now tells us that $\beta\rho > 1$ and the λ-leximin path is Nirvana for all λ. Note that the formula for $\frac{u^2}{u^1}$ given in equation (12.7), which is also the formula for $\frac{u^{t+1}}{u^t}$ on any Nirvana path, says that on Nirvana paths, utility increases without bound if and only if $\beta\rho > (1+\gamma\lambda)^{\theta}$. This is surely true if $\beta\rho > 1$, which proves statement B. □

We note, from the last part of the proof, that there may be Nirvana paths at which utilities increase without bound but the forest goes to zero size: the converse of statement B is not generally true. This happens when we have Nirvana with $1 > \beta\rho$. Nirvana implies that $\lambda > \lambda^*$ which is equivalent to $\beta\rho > (1 + \hat{\gamma}\lambda)^\theta$.

Let us try some parameter values: I suggest

$a = 0.75$	(labour's share is about 75 per cent)
$c = 0.5$	(people value the environment and commodities equally)
$\hat{\gamma} = 13$	(productivity growth at 2 per cent annum implies that over a generation of 25 years, the growth factor is 1.64. Assuming employment in R&D is 5 per cent of the labour force, this gives $.05\gamma = 0.64$)
$\rho = 1.64$	(assume a 2 per cent per annum rate of regeneration).

With these values, we compute that $\lambda^* = -.02$ and so we have Nirvana at all λ.

If we think of the biosphere as necessary for health, then c should be quite large. If a and ρ are given the above values, then $\lambda^* \leq 0$ as long as $c \geq 0.32$. This seems fairly optimistic.

Suppose, then, that $c = 0.333$. Then with a and $\hat{\gamma}$ as above, the overall leximin path is given by $\lambda = .55$, a far cry from what we see. This suggests that a weakness in the model is the assumption that the rate of technical change is linear in labour: in reality, there must be fairly strongly decreasing returns to R&D in labour, due to the high level of education needed to be productive in that sector. Perhaps a more realistic model would postulate a limited supply of labour capable of working in the R&D sector. Endogenizing *that* supply would require putting an education sector into the model.

12.6 Concluding contemplations

Let me conclude with some conjectures and remarks.

1. Perhaps the most important conclusion is that the leximin social welfare function does not automatically relegate us to a world with no human progress, that is, no increase in well-being over time. Clearly this can be the case with exogenous technical change: but it holds even if technical change is endogenous and costly. If the technology that transforms labour into technical progress is sufficiently productive (the value of k in theorem 5), then leximin appears to imply increasing well-being over time – at least on the domain of constant growth paths.

   The most general formulation of the conditions under which leximin involves an increase in welfare over time is given in Silvestre (2002).

   I believe this is a potentially important observation. Intergenerational maximin is often associated with the view that utilities must be constant over time, and hence, in the welfare sense, there is no progress. Because many consider this to be an unacceptable outcome (is it really?), the maximin or leximin social welfare function is not taken as seriously as it should be. We may, however, live in a world where the conditions hold where intergenerational leximin implies increasing utilities over time.

2. An important assumption, which I believe is responsible for the somewhat pessimistic results in the Cobb-Douglas case, is that technological progress occurs in the housing industry but not in the hiking industry. This, I think, is a realistic assumption. The attraction that Nature has for us is, I believe, that we enjoy it unmediated by sophisticated technology. If we interpret 'nature' as the services of the biosphere that sustain life, then the appropriate assumption is that our capacity to create consumption commodities using the biosphere as a resource improves more rapidly than our capacity to create health from the biosphere.
3. Other sentient beings. Our discussion does not take into account the welfare of other sentient beings that use the natural resource. Of course, doing so could radically alter our conclusions about forest use. The 'green' definition of sustainability is perhaps a quick reduced form for modelling the welfare of animals who also use the forest for life. Even if, today, our society will not reach a consensus to include the welfare of other sentient beings in the calculus of equity, we might wish to retain the flexibility on this matter.

   Our possible concern with animals is one which would cause us to alter the arguments of the intergenerational utility function, to include more than one representative agent at each date. Indeed, we may wish to represent different agents as being of different sizes.
4. If we relax the assumption of the single representative agent at each date, there are plausible ways of doing so short of admitting animals as citizens. We should recognize that there are different types of humans, in particular, humans who have different wealth and income levels. The equality-of-opportunity approach directs us to consider the case of types of people who live simultaneously and who have different welfare levels because of circumstances beyond their control – for instance, because they were born in different countries, or in different families in the same country, or with different genetic dispositions in the same family. All these differences among people are circumstances, and arguably, an equal-opportunity ethic would declare that it is unjust that these differences should entail different welfare levels for the individuals in question.

   Thus, the intergenerational equity problem would here be to maximize the welfare level of the worst-off person who ever lives, assuming that we do not complicate the model further by allowing people to expend different degrees of voluntary effort. It will now be appropriate to have intragenerational transfers through taxation, and so the optimization problem is more difficult: there are at least two controls, the use of the natural resource and intragenerational taxation. As well, if there are poor and rich families, we probably should take into account the process of formation of human capital, and hence investment will take the form not only of building houses, but of education.

Roberto Veneziani and I (2004) have recently studied a version of this problem, but without the natural resource: we are concerned only with the nature of educational investment and taxation to make the worst-off person who ever lives as well off as possible, in a society with two dynasties, distinguished by their initial endowments of human capital. Discussing this problem is beyond my scope here. Let me just offer one remark. Our attention is naturally focused, in such a problem, on the poor, low human-capital people at each date: they are the ones who are the worst-off. Because demand for hiking in the forest is a normal good, it will be optimal to consume more of the forest as investment than if everyone were well off.

This is, perhaps, one of the most important ways that our model fails to capture what is ethically important in today's world. Those at the conference where this chapter was presented have demands for preservation of the natural environment that are associated with having incomes in the top one per cent or so of world income. But intergenerational equity, of the equal-opportunity kind that I have been discussing, probably will focus upon the welfare of those who live in South Asia and Africa, and perhaps Latin America: they are the worst-off today, and will be for some time to come. It may be that the optimal solution entails running down the natural environment until a point that technological advance has become sufficient that it is no longer necessary to do so; but perhaps this path can be avoided with sufficient transfers from the north to the south.

The problem is interesting, because, to some extent at least, the natural environment is a commons, from which all countries can harvest timber. If the citizens of China or Brazil or the Sudan are close to their minimal consumption of housing, then one can only expect them to consume from the natural commons at a rate faster than citizens of the rich countries would like. The travesty is not the overconsumption of the poor countries *per se*, but rather the refusal of the advanced countries to transfer more resources to them to substitute for timber. Such transfers would increase the welfare of the worst-off while allowing them to deplete the natural resource at a slower rate. It is difficult to fault the poor countries for their use of the forest given the constraints that they face, and an equal-opportunity intergenerational ethic.

The equal-opportunity ethic, as I have described it here, ignores national boundaries. This is a contentious move, and is an instance of what is currently called the cosmopolitan view. A number of political philosophers who are quite egalitarian in the context of the single nation, are not so when it comes to the international community. (See, for example, Rawls (1999) and Nagel (2005).[4]) My own view is that, in the next millennium, when people look back on our time, they will find the most egregious inequalities to be those in income per capita across nations. When even the most humanitarian countries transfer only about one per cent of their national income to other countries in aid, we may say that human beings have scarcely begun to view themselves as citizens of a world, rather than of a nation-state.

The most challenging problem, then, is to study the issue of sustainability when we address not only intergenerational equity, but intragenerational and international equity as well.

Notes

1. Note that the model of the previous sections is a special case of this one, where we take $a = 0$, $b = c$, and all labour is assigned to the R&D industry, generating a rate of growth of $\alpha(t)$ equal to $\hat{\gamma}$.
2. The reader may check that Theorem 1 is a special case of Theorem 5, where we let $a = 0$ and $\lambda = 1$.
3. Of course, *pace* section 1, the 'overall' leximin path is not the *true* universal leximin path which may well require varying the labour employed over time in the two industries.
4. Nagel (2005) advocates the view that the proper relations among nations at this time are ones governed by bargaining, not by justice. If and when a supranational state comes into being, then justice would become the appropriate currency.

References

Lucas, R.E. (1988) 'On the Mechanism of Economic Development', *Journal of Monetary Economics*, vol. 22, pp. 3–42.

Nagel, T. (2005) 'The Problem of Global Justice', *Philosophy & Public Affairs*, vol. 33, pp. 113–47.

Rawls, J. (1999) *The Law of Peoples* (Cambridge, MA: Harvard University Press).

Roemer, J.E. and R. Veneziani (2004) 'What We Owe Our Children, They Their Children, . . .', *Journal of Public Economic Theory*, vol. 6, pp. 637–54.

Silvestre, J. (2002) 'Progress and Conservation under Rawls's Maximin Principle', *Social Choice and Welfare*, vol. 19, pp. 1–27.

Uzawa, H. (1965) 'Optimum Technical Change in an Aggregative Model of Economic Growth', *International Economic Review*, vol. 6, pp. 18–31.

13

Intergenerational Justice, International Relations and Sustainability*

John E. Roemer and Roberto Veneziani

13.1 Introduction

During the next thirty to fifty years, the populous, poor countries of China and India will, one hopes, approach the current per capita income of countries in the highly developed world today. In purchasing-power-parity terms, China's per capita in 2003 was $4,990, while Britain's was $27,650.[1] At an annual rate of real growth of 7% per annum – considerably less than China's growth rate over the past twenty years – China's per capita income will overtake Britain's current income in twenty-five years' time. One can expect that Chinese and Indian citizens will desire at least the standard of living that Britons enjoyed in 2003. The pressure on the natural resource base of the earth, in particular the biosphere, will in all likelihood be immense.

In this chapter, we model a world with two 'countries', called the North and the South. The North is advanced in possessing a technology that can convert the world's natural resource – which we call the 'forest' – into commodities that consumers enjoy – which we call 'houses' – at a high rate. (That is, relatively little of the forest must be harvested to produce a house.) In contrast, the South has a less efficient technology, requiring more forest harvesting per house produced. The forest is a global commons: thus, without further restrictions, both North and South can harvest the forest as they wish. Finally, the utility of citizen-consumers in both countries is a function of houses consumed and of the size of the unharvested forest. One may think of the unharvested forest as the unpolluted biosphere, or, the rate of 'biospheric services' delivered to people, a (global) public good.

* We are indebted to Marc Fleurbaey, Toshihiro Ihori, Marco Mariotti, Ngo Van Long, and participants in seminars at the Catholic University of Milan, Queen Mary, the State University of Milan, and to the IEA conference on intergenerational equity, for comments and suggestions. The usual disclaimer applies.

We postulate that the world just described goes on for many generations. Technological advance occurs over time, in both North and South. At each date, there are two agents – a representative of her generation in the North and in the South. (These two agents, however, may be of different sizes.) We further postulate that each country aims to achieve intergenerational justice for its citizens: it desires to consume the forest over time in that way which maximizes an intergenerational social welfare function. Relations between the two countries, however, may take on a variety of forms: the countries may behave strategically and selfishly in a game in which their strategies are intertemporal sequences of national forest consumption; they may bargain with each other, entering into treaties concerning forest consumption over time; or they may collectively adopt the cosmopolitan view, and cooperate to choose a path of forest consumption which maximizes *world* social welfare, ignoring national boundaries.

Sustainability has at least two meanings in this context. One, the anthropocentric one, defines a path of forest consumption as sustainable if the welfares of the human subjects are non-decreasing over time. The second, green definition defines a path of forest consumption as sustainable if the forest does not disappear asymptotically, or, more strongly, (eventually) does not decrease in size over time. We will be interested in sustainability in both senses.

We next indicate our approach to intertemporal (national) and international justice. We adopt the equal-opportunity (EO) approach as far as intertemporal justice is concerned. The only differences among the representative generational agents of a given nation are due to the time of their births: these are the technology they enjoy, the size of the forest, and the nature of the other country. These differences are all beyond the control of these agents, and so equality of opportunity implies that justice requires equalizing their utilities: to be somewhat more precise, lexical maximization of minimal utilities over time, or leximin. Thus, we assume that the social preference order for each country, in either the non-cooperative or bargaining cases described above, is the leximin preference order over utilities of all future generations.

There are several contemporary views in political philosophy about what constitutes international justice. John Rawls and Thomas Nagel argue, in different ways, that such justice does not require the maximization of a joint social welfare function: rather, it requires bargaining among countries. The cosmopolitan view, advocated by, for example, Thomas Pogge, asserts that national boundaries are morally arbitrary, and that the same principles should be applied to international as to intranational justice. Our task is not to adjudicate this debate here: we shall study both the bargaining and the cosmopolitan solutions.

The chapter proceeds by defining and studying, first, the non-cooperative game between the North and the South, where each country chooses a strategy, consisting of a path of forest consumption, to leximin the stream of welfares among the generations in its country. Secondly, we study the cosmopolitan solution; thirdly, we study a bargaining solution. Note that it may be in the interest of the North to transfer houses (commodities), gratis, to the South, because the North can produce those houses with less damage to the forest than can the South. Such transfers will be important in both the bargaining and cosmopolitan solutions.

We will be interested in characterizing the streams of welfare of the two dynasties, the nature of international transfers, and the evolution of the forest in the three solutions. Doing this in all generality is a task too large for one chapter, and so we will make simplifying assumptions as we need them to achieve simple results.

13.2 The dynamic environment

We model the problem in a stark way. There are two societies that exist for an infinite number of generations. In each country, agents are identical and population size is constant; a fraction f of the world population lives in country 1, while a fraction $1 - f$ lives in country 2. We assume that there is a natural environment, to be thought of as a resource that can either be used to produce consumption goods, or can be enjoyed in its pristine state. Think of nature as a forest, which can either be used for timber, to build houses, or for hiking and recreation.[2] The forest has a natural rate of growth, or regeneration, and let us assume that it is depleted only when used for timber, but not for hiking. Clearly, there is a rate at which wood can be harvested, so that the forest would remain of constant size. One might think of this rate as the *sustainable* rate of harvesting the forest, but that rate is not of any particular ethical significance, because we have not yet related it in any way to human welfare.

More formally, let R^t denote the stock of natural resources (the forest) at t, with R^0, the initial size of the forest, given. Let ρ denote the natural rate of growth of R; and let I_j^t denote the harvest of country j, $j = 1, 2$, per world capita, at date t. Then,[3]

$$R^t + I_1^t + I_2^t = \rho R^{t-1}. \tag{13.1}$$

Let $\Re_+$ denote the nonnegative real numbers: we suppose that there exist two strictly increasing functions $h_1 \colon \Re_+ \to \Re_+$ and $h_2 \colon \Re_+ \to \Re_+$ such that if an amount of forest I_j^t is harvested at date t in country j, $j = 1, 2$, then an amount of housing $H_j^t = h_j(I_j^t)$ is produced; differences in h_1 and h_2 reflect the different development stages of the two economies. To be specific, suppose that members of both societies strive to improve their lives, and to this end, they engage in research which produces technological improvements – that is, at least under normal conditions, there will be a rate of productivity increase in the use of timber to make houses. We assume that at the beginning date $t=1$, country 1 is advanced, while country 2 is backward (poor), and that technical knowledge cannot be freely transferred across borders: for the sake of simplicity, we assume that the cost of transferring knowledge is infinite. Instead, within each country, each generation costlessly passes down its technology to the next one. Thus, technological striving produces a positive externality for future generations living in the same country.

More formally, we assume that there exist two exogenous, strictly increasing sequences $\{\alpha^t | t=1,\ldots\}$ and $\{\beta^t | t=1,\ldots\}$, describing technical knowledge at date t in countries 1 and 2, respectively.[4] If I_1^t and I_2^t are the harvest of the forest as timber at date t, in countries 1 and 2, respectively, we suppose that the per capita production of housing is $\alpha^t I_1^t$ and $\beta^t I_2^t$, respectively, where $\alpha^t \geq \beta^t$, all t.

Let $\Re_+^2 = \Re_+ \times \Re_+$. We assume that agents in the two societies have identical preferences, which can be described at each t by a strictly increasing, twice differentiable, and strictly quasi-concave function u: $\Re_+^2 \to \Re_+$ of consumption goods produced from the part of the forest harvested as timber at t, H^t, and of the hiking in the pristine forest that remains uncut; that is, the recreational use of the forest, which we assume to be just measured by the size of the forest left unharvested. Thus, the utility of an agent living in country j at date t is $u_j^t = u(R^t, H_j^t)$.[5] Let $u^{\min} = \min_{R,H} u(R, H)$; for the sake of simplicity, we assume that $u(0, H) = u(R, 0) = u^{\min}$, for any R, H.[6]

It is important to note that the welfare of an individual in each society is here a function only of her own consumption, not of the consumption of her descendants. You can think of this assumption in one of two ways: either individuals indeed *do not* worry about their descendants, or, more persuasively, the kind of well-being with which we, the impartial ethical observer, should be concerned, is the *standard of living* of individuals, rather than the kind of happiness they get when contemplating their children's and grandchildren's lives, which is properly called *welfare*. *We*, the ethical observer, will take care of the children and grandchildren.

Since the forest is modelled here as a public good with no congestion, the two countries face a global commons problem and we can assume that at each t, the advanced country may be willing to transfer a fraction τ^t of its output $\alpha^t I_1^t$ to the poor country to mitigate the problem. However, consistent with the assumption that technical knowledge cannot be costlessly transferred across borders, we assume that the international transfer of output, too, is costly. To be specific, we assume that there is a decreasing, twice-differentiable function π: $[0, 1] \to [0, 1]$, with $\pi(0) = 1$ and $\pi(1) \geq 0$, such that if the advanced country gives up an amount of output $\tau^t \alpha^t I_1^t$, the poor country only receives $\pi(\tau^t)\tau^t \alpha^t I_1^t$. Thus, if international transfers are possible, the utility of a country 1 citizen at date t is $u[R^t, (1-\tau^t)\alpha^t I_1^t / f]$ while the utility of a country 2 citizen at date t is $u[R^t, (\beta^t I_2^t + \pi(\tau^t)\tau^t \alpha^t I_1^t)/(1-f)]$.

13.3 The non-cooperative model

In this section, we study the model where countries do not cooperate and there is no international transfer of resources, so that at date t the utility of a country 1 citizen is $u(R^t,\ \alpha^t I_1^t/f)$, while the utility of a country 2 citizen is $u(R^t,\ \beta^t I_2^t/(1-f))$. We are interested in the allocation of use of the natural resource in each country – where by 'use' we mean its use in the production of commodities (houses) – and we may denote an allocation as a pair $I^t = (I_1^t, I_2^t)$ which specifies the amount that will be harvested and used for the production of houses at t in each country; and thus the size of the forest at date t that remains in its uncut state, R^t.

Let $\overline{R} = \{R^t | t = 1, \ldots\}$ and $\overline{I}_j = \{I_j^t | t = 1, \ldots\}$, $j = 1,\ 2$, denote, respectively, generic paths of the natural resource and of the harvest in country j. Since the two countries behave non-cooperatively, we assume that intergenerational equality of opportunity requires each country j to find the path $\overline{I}_j$ that maximizes the minimum level of welfare across all generations, given $\overline{I}_i$, $i \neq j$. In other words, for country j, we wish to solve the *non-cooperative maximin programme* (MP_1):[7]

$$
(MP_1) \quad \begin{aligned}
&\max_{I_1^t} \min_{t \geq 1} u(R^t, \alpha^t I_1^t/f),\\
&\text{subject to: } R^t = \rho R^{t-1} - I_1^t - I_2^t,\ \text{all } t \geq 1,\\
&I_1^t, R^t \geq 0,\ \text{all } t \geq 1, \text{ and given } \overline{I}_2 \text{ and } R^0.
\end{aligned}
$$

The *non-cooperative maximin programme* for country 2 (MP_2) is defined similarly. There may, indeed, be many maximin solutions: what we really are interested in is the *lexicographic minimum solution*, which we define as follows. Let $\mathbf{P}_{\mathrm{jN}}(1)$ denote the maximin programme (MP_j), $j = 1,\ 2$. Consider the sequence of programmes:

$$
\mathbf{P}_{1\mathrm{N}}(n) \quad \begin{aligned}
&\max_{I_1^t} \min_{t \geq n} u(R^t, \alpha^t I_1^t/f)\\
&\text{subject to: } R^t = \rho R^{t-1} - I_1^t - I_2^t,\ \text{all } t \geq n,\\
&I_1^t, R^t \geq 0,\ \text{all } t \geq n;\ \text{given } I_2^t,\ \text{all } t \geq n, \text{ and } R^{n-1},
\end{aligned}
$$

for country 1. The sequence of programmes $\mathbf{P}_{2\mathrm{N}}(n)$ for country 2 is similarly defined.

In other words, if $\overline{I}_j$ solves $\mathbf{P}_{\mathrm{jN}}(1)$, it is a maximin solution for country j. If it solves $\mathbf{P}_{\mathrm{jN}}(2)$, it is, among all maximin solutions, one which is also maximin for generations two and beyond in country j; and so on. Then, we can define a *leximin solution* for country j as follows.

Definition 1. A path $\bar{I}_j$ is a *leximin solution for country j* (LS_j) *given* $\bar{I}_i$, $i \neq j$, if it solves $\mathbf{P_{jN}}(n)$, all $n \geq 1$.

Next, we can define a *global non-cooperative leximin solution*.[8]

Definition 2. A pair $\bar{I}_1$, $\bar{I}_2$ is a *global non-cooperative leximin solution* (NCLS), if $\bar{I}_1$ is a LS_1 given $\bar{I}_2$, and $\bar{I}_2$ is a LS_2 given $\bar{I}_1$.

A remark is in order here: there are various ways of modelling the non-cooperative interaction between the two countries and some would argue that the asymmetries in economic and political weight between countries in the international arena often imply that treaties are signed and implemented that are promoted by the advanced countries to foster their interests – which may include equity concerns towards future co-nationals, as in our model. Albeit relevant, these issues need not concern us here, given our focus on countries such as China and India, which have already acquired a significant weight in the international arena. In this setting, we believe that it is reasonable to assume the North and the South to be relatively equal in staying power, and adopt the NCLS as the appropriate non-cooperative solution concept, rather than an asymmetric Stackelberg-type concept.

We can now characterize the path of utilities u_j^t, $j = 1, 2$, at a NCLS.

Proposition 1. *At a NCLS,* $u_j^t \leq u_j^{t+1}$, *all* $t \geq 1$, $j = 1, 2$.

Proof. At a LS_j, $\rho R^{t-1} > I_j^t$, all t, j, and thus at a NCLS $I_j^t > 0$, all t, j. Then, consider country j. Suppose that $u_j^t > u_j^{t+1}$, some t: by slightly decreasing I_j^t, it is possible to increase R^t, and all utilities after t, without lowering u_j^t below u_j^{t+1}, contradicting the assumption that the original solution was a LS_j. *QED.*

Next, we say that a generation in country j is unconstrained if the prescribed path at date t is just the value of the forest that agent t would decide upon were she to optimize selfishly, given her endowment and taking the other country's choices parametrically, and ignoring future generations.

Definition 3. Generation t in country j is *unconstrained on* $\bar{I}_j$ if I_j^t is the individually optimal value for t, given R^{t-1} and taking I_i^t, $i \neq j$, parametrically. If it is not unconstrained on $\bar{I}_j$, then we say that generation t in country j is *constrained.*

Let $I_j^{t^*}$ denote the individually optimal value for agent t in country j.

Proposition 2. *If* $\bar{I}_j$ *is a* LS_j *and agent t in country j is constrained, then* $I_j^t < I_j^{t^*}$.

Proof. Suppose not. Then I_j^t could be decreased to $I_j^{t^*}$, thus increasing R^t and the welfare of all future generations, a contradiction. *QED.*

Given Proposition 2, the following result can be proved.

Proposition 3. *At a LS_j, if $u_j^t < u_j^{t+1}$, some t, then agent t is unconstrained.*

Proof. Suppose generation t in country j were constrained. Then by Proposition 2, she could increase u_j^t by increasing I_j^t a small amount. By Proposition 1 this raises the value of $\mathbf{P_{jN}}(t)$, and hence the original solution was not leximin. *QED.*

Then, Lemmas 1 and 2 derive sufficient conditions for a path $\bar{I}_j$ to be the LS_j.[9]

Lemma 1. *For a given $\bar{I}_i$ such that $I_i^t = z_i R^{t-1} + w_i^t$, all t, with $r_i = (\rho - z_i) > 1$, let $\bar{I}_j$ be such that $u_j^t = u_j^{t+1}, 0 < I_j^t < I_j^{t^*}$, and $R^t \leq R^{t-1}$, all t. Then $\bar{I}_j$ is the LS_j.*

Proof:

1. Suppose not. Then, there is another feasible $\bar{I}_j'$ such that $u_j'^t \geq u_j^t$, all t, with strict inequality some t. Let u_R and u_H denote, respectively, the derivative of u with respect to the first and the second argument. Let $u_R^t = u_R(R^t, H_j^t)$ and $u_H^t = u_H(R^t, H_j^t)$ be the values of the derivatives at t in the proposed solution.
2. Consider country 1 (a similar argument applies to country 2). Since $I_1^t > 0$, the choice $(R^t + \varepsilon, I_1^t - \varepsilon)$ is feasible for small $\varepsilon > 0$. Viewing u^t as a function of (R^t, I_1^t), its directional derivative in the direction $(1, -1)$ must be negative, by strict quasi-concavity, for if it were positive, we could increase the value of u_1^t while not decreasing any other utility on the path. Therefore $u_R^t < (\alpha^t/f)u_H^t$, all t, while $u_1'^t \geq u_1^t$, all t, implies $u_R^t\ [R'^t - R^t] + (\alpha^t/f)u_H^t$ $[I_1'^t - I_1^t] \geq 0$, all t, with strict inequality unless $(R'^t, I_1'^t) = (R^t, I_1^t)$.
3. Let t^* be the first generation for which $(R'^t, I_1'^t) \neq (R^t, I_1^t)$. Without loss of generality, we may assume $t^* = 1$. Then $I_1'^1 - I_1^1 = c^1 > 0$. Let $I_1'^2 - I_1^2 = c^2$; then $R'^2 - R^2 = -r_i c^1 - c^2$. Since $u_1'^2 \geq u_1^2$, $-u_R^2\ [r_i c^1] + [(\alpha^2/f)u_H^2 - u_R^2]$ $c^2 \geq 0$, which implies $c^2 \geq 0$, using the inequality established in step 2. Iterating over t, we have that $u_1'^t \geq u_1^t$, implies $-u_R^t\ [r_i c^{t-1} + (r_i)^2 c^{t-2} + \cdots + (r_i)^{t-1} c^1] + [(\alpha^t/f)u_H^t - u_R^t]\ c^t \geq 0$, which implies that $c^t \geq 0$, all t.
4. Note that $R'^t - R^t = -\sum_{\tau=1}^{t-1} (r_i)^{t-\tau} c^\tau$. Since $c^1 > 0$, $r_i > 1$, and $c^t \geq 0$, all t, we have $\lim_{t\to\infty} R'^t - R^t = -\infty$. But $\{R^t\}$ is bounded above by R^0, by our premise. It follows that eventually $R'^t < 0$, an impossibility which establishes the lemma. *QED.*

Finally, Lemma 2 focuses on unconstrained paths.

Lemma 2. *Let $\bar{I}_j$ be a path such that, at all t, $u_j^t \leq u_j^{t+1}$ and each agent is unconstrained. Then $\bar{I}_j$ is the LS_j.*

Proof. u_j^1 cannot possibly be greater, and so $\bar{I}_j$ solves $\mathbf{P_{jN}}(1)$, $j=1, 2$. Hence R^1* is determined. By induction, the path is the leximin path. *QED*.

Since we are interested in analyzing environmental, intergenerational, and international issues under different institutional assumptions, our next step is to put more structure on the utility function, in order to characterize more precisely the dynamics of welfare and natural resources. In particular, we shall assume that agents have a Cobb–Douglas utility function: this is a standard assumption which is made here mainly for analytical convenience.[10]

Assumption 1. (A1): *Let* $u(R, H)=R^bH^{1-b}$, *some* $0<b<1$.

Under (A1), we compute that if both countries optimize selfishly then

$$I_1^t = (1-b)(\rho R^{t-1} - I_2^t), \quad \text{all } t, \tag{13.2}$$

$$I_2^t = (1-b)(\rho R^{t-1} - I_1^t), \quad \text{all } t. \tag{13.3}$$

Therefore if their expectations are realized at all t, from (13.2)–(13.3) it follows that

$$I_j^t = \left(\frac{1-b}{2-b}\right)\rho R^{t-1}, \quad \text{all } t, j=1,2, \tag{13.4}$$

which in turn implies

$$R^t = \frac{b\rho}{2-b}R^{t-1}, \quad \text{all } t. \tag{13.5}$$

The resulting indirect utility functions are

$$u_1^t = b^b(1-b)^{1-b}\left(\frac{\alpha^t}{f}\right)^{1-b}\frac{\rho R^{t-1}}{2-b}, \quad \text{all } t, \tag{13.6}$$

$$u_2^t = b^b(1-b)^{1-b}\left(\frac{\beta^t}{1-f}\right)^{1-b}\frac{\rho R^{t-1}}{2-b}, \quad \text{all } t, \tag{13.7}$$

which imply

$$\frac{u_j^{t+1}}{u_j^t} = (\gamma_j^{t+1})^{1-b}\frac{b\rho}{2-b}, \text{ all } t, j=1,2. \tag{13.8}$$

Theorem 1. *Let* $\alpha^t/\beta^t=\lambda>1$, $\alpha^{t+1}/\alpha^t=\gamma>1$, *all* t. *Under (A1), if* $\rho \geq \rho_{\min}^{NCLS} = \frac{2-b}{b}\gamma^{b-1}$, *then the unconstrained NCLS is recursively defined by (13.4). Moreover, if* $\rho=\rho_{\min}^{NCLS}$ *then* $u_j^t=u_j^{t+1}$, *all* t, $j=1, 2$, *and* R^t *decreases to zero; while* *if $\rho > \rho_{\min}^{NCLS}$, then $u_j^t < u_j^{t+1}$, all t, j = 1, 2, and R^t decreases to zero if $\rho < (2-b)/b$, remains constant if $\rho = (2-b)/b$, and increases to infinity if $\rho > (2-b)/b$.*

Proof:

1. Given $\alpha^t/\beta^t = \lambda$ and $\alpha^{t+1}/\alpha^t = \gamma$, all t, it follows from (13.8) that selfish optimization generates $u_j^t = u_j^{t+1}$, all t, $j = 1, 2$, if and only if $b\rho(\gamma)^{(1-b)} = (2-b)$ and $u_j^t < u_j^{t+1}$, all t, $j = 1, 2$, if and only if $b\rho(\gamma)^{(1-b)} > (2-b)$. In either case, by Lemma 2 this is a NCLS.
2. The results concerning the dynamics of R^t follow from $\gamma > 1$ and (13.5). *QED.*

Theorem 1 tells us that, if structural differences between countries remain constant over time, sustainability in the literal sense of a non-shrinking forest size occurs if and only if love of the forest, b, is sufficiently great. Moreover, the value of ρ at which the North and the South become unconstrained is the same.

Theorem 2 characterizes economies with $\rho < \rho_{\min}^{NCLS}$ and constrained countries.

Theorem 2. *Let $\alpha^t/\beta^t = \lambda > 1$, $\alpha^{t+1}/\alpha^t = \gamma > 1$, all t. Under (A1), if $b\gamma^{1-b} < 1$ and $(1 - b + b\gamma^{1-b}/b\gamma^{1-b}) \le \rho < \rho_{\min}^{NCLS}$, then $I_1^t = x(\rho\gamma^{1-b} - 1)\gamma^{b-1}R^{t-1}$, $I_2^t = (1-x)(\rho\gamma^{1-b} - 1)\gamma^{b-1}R^{t-1}$, $R^t = \gamma^{-(1-b)}R^{t-1}$, all t, is a NCLS for any $x \in (x_1, 1-x_1)$ where $x_1 = \frac{b\rho\gamma^{1-b} - 1}{b(\rho\gamma^{1-b} - 1)}$; every generation is constrained; $u_j^t = u_j^{t+1}$, all t, j = 1, 2; and R^t decreases to zero.*

Proof:

1. Compute that the suggested solution implies $u_j^t = u_j^{t+1}$, all t, $j = 1, 2$ and $R^t > R^{t+1}$, all t. Next, note that $\rho - x(\rho\gamma^{1-b} - 1)\gamma^{b-1} > 1$ and $\rho - (1-x)(\rho\gamma^{1-b} - 1)\gamma^{b-1} > 1$ all $x \in (x_1, 1-x_1)$. Finally, since $1 < b\rho(\gamma)^{1-b}$ then $1 > x > 0$, all $x \in (x_1, 1-x_1)$.
2. Consider country 2: if $x > x_1$ then I_2^t is less than the individually optimal level, all t, given $\bar{I}_1$. Therefore, by Lemma 1, $\bar{I}_2$ is the LS_2. Similarly, if $x < 1 - x_1$, then I_1^t is less than the individually optimal level, all t, given $\bar{I}_2$. Hence by Lemma 1, $\bar{I}_1$ is the LS_1. Finally, since $\gamma^{b-1} < 1$, then $\lim_{t\to\infty} R^t = 0$. *QED.*

13.4 Pareto efficiency in the one-period model

Our next task is to introduce the possibility of transfers of houses from the North to the South. Given our interest in the unconstrained path, we first characterize the Pareto efficient allocations in the one-period model. Thus, we drop all time subscripts and note that the set of Pareto efficient allocations is the solution of

$$\text{(PA)}\quad \begin{aligned} &\max_{I_1,I_2,\tau} u\left(\rho R^0 - I_1 - I_2, \frac{(1-\tau)\alpha I_1}{f}\right), \\ &\text{subject to: } u\left(\rho R^0 - I_1 - I_2, \frac{\beta I_2 + \pi(\tau)\tau\alpha I_1}{1-f}\right) \geq k, \\ &I_1, I_2 \geq 0, \rho R^0 \geq I_1 + I_2, \text{ and } \tau \in [0,1]; \text{ and given } R^0. \end{aligned}$$

Let ϕ denote the Lagrange multiplier of PA, let u_R^N denote the derivative of the utility of the North with respect to the first argument and likewise for u_H^N, u_R^S, u_H^S. In order to simplify the notation, let $\sigma(\tau) = \pi(\tau)\tau\alpha$ with $\sigma'(\tau) = \alpha[\pi(\tau) + \pi'(\tau)\tau]$. The following first order conditions can be derived:

$$I_1 : -u_R^N + \frac{(1-\tau)\alpha}{f} u_2^N - \phi u_R^S + \phi \frac{\sigma(\tau)}{1-f} u_H^S = 0, \tag{13.9}$$

$$I_2 : -u_R^N - \phi u_R^S + \phi \frac{\beta}{1-f} u_H^S = 0, \tag{13.10}$$

$$\tau : -\frac{\alpha}{f} I_1 u_H^N + \phi \frac{\sigma'(\tau)}{1-f} I_1 u_H^S = 0, \tag{13.11}$$

$$\phi : u^S = k. \tag{13.12}$$

From (13.9) and (13.10) it follows that

$$\phi = \frac{\alpha(1-\tau)}{f} \frac{1-f}{\beta - \sigma(\tau)} \frac{u_H^N}{u_H^S}, \tag{13.13}$$

and substituting the latter expression into (13.11) we have

$$\sigma'(\tau) = \frac{\beta - \sigma(\tau)}{1-\tau}, \tag{13.14}$$

or equivalently, using the definition of $\sigma(\tau)$

$$1/\lambda = (1-\tau)\tau\pi'(\tau) + \pi(\tau). \tag{13.15}$$

If $\tau\pi''(\tau) < 2|\pi'(\tau)|$, all $\tau \in [0, 1]$, – i.e., if either π is concave or second order effects are not too significant – (13.15) identifies a unique value of the transfer rate at all Pareto optimal allocations. Given the assumptions on π, $\tau \in (0, 1)$; while if λ tends to one τ tends to zero and as λ grows indefinitely large τ tends to one. This is not surprising given that the only reason for country 1 to transfer part of its output to country 2 is the technological gap. Next, using (13.13) and (13.14) into (13.10),

$$\frac{\alpha(1-f)}{\sigma'(\tau)f}\frac{u_H^N}{u_H^S}\frac{\beta}{1-f}u_H^S = u_R^N + \frac{\alpha(1-f)}{\sigma'(\tau)f}\frac{u_H^N}{u_H^S}u_R^S,$$

or, equivalently,

$$\beta = \frac{\sigma'(\tau)}{\alpha}f\frac{u_R^N}{u_H^N} + (1-f)\frac{u_R^S}{u_H^S}. \tag{13.16}$$

Equation (13.16) is more transparent if we note that u_R^N/u_H^N is just the marginal rate of substitution between housing and consumption of the forest for the North, and likewise for u_R^S/u_H^S. Thus, let H_1 and H_2 denote housing consumption for the North and for the South, respectively. We can derive the relation between H_1, H_2, and R – the part of the forest left uncut – as follows. By definition

$$H_2 = \frac{\beta I_2 + \pi(\tau)\tau\alpha I_1}{1-f},$$

which, given $I_1 = \frac{fH_1}{\alpha(1-\tau)}$ and $I_2 = \rho R^0 - R - I_1$, can be written as

$$H_2 = \frac{\beta(\rho R^0 - R)}{1-f} + \frac{\sigma(\tau)-\beta}{1-f}\frac{fH_1}{\alpha(1-\tau)},$$

or, equivalently,

$$(1-f)H_2 + \frac{\sigma'(\tau)}{\alpha}fH_1 = \beta(\rho R^0 - R). \tag{13.17}$$

We can now derive our main result concerning Pareto-optimal allocations.

Theorem 3. *Let u(R, H) be any quasi-concave, differentiable utility function. Let K be a positive constant. In the one-period model:*

A. *If* $\tau\pi''(\tau) < 2|\pi'(\tau)|$, *all* $\tau \in [0, 1]$, *then the transfer* τ *is invariant across all Pareto-efficient allocations and is determined by* (13.15).
B. *If* $u_R/u_H = q(R)H^\mu + Kq(R)$, $\mu = 0, 1$, *where q is a strictly monotone decreasing function, then* $I = I_1 + I_2$ *is invariant across all Pareto-efficient allocations.*

Proof:
Part A. It follows from the first order conditions of (PA), in particular (13.15).
Part B. By Part A, in order to characterize the set of Pareto-efficient allocations we need to focus only on (13.12), (13.16), and (13.17). Suppose first

that $\mu = 0$ and the marginal rate of substitution is $u_1/u_2 = q(R)\ (1+K)$. Then, by (13.16) it follows

$$\frac{\beta}{fq(R)(1+K)} = \left[\frac{\sigma'(\tau)}{\alpha} + \frac{1-f}{f}\right], \tag{13.18}$$

which determines R since q is strictly monotonic.

Next, suppose that $\mu = 1$ so that $u_R/u_H = q(R)\ (H+K)$. By (13.16) it follows that

$$\beta = \frac{\sigma'(\tau)}{\alpha} fq(R)(H_1+K) + (1-f)q(R)(H_2+K),$$

and combining the latter expression with (13.17), it follows that

$$\beta(\rho R^0 - R) = \frac{\beta}{q(R)} - K\left[1 - f + \frac{\sigma'(\tau)}{\alpha} f\right], \tag{13.19}$$

which determines R since q is strictly monotonic. *QED.*

Theorem 3 is fairly general. Part A applies to all utility functions, while the class of functions in Part B is quite large and includes quasilinear utility, Cobb-Douglas, and Stone–Geary, among others. However, it is worth noting that intergenerational Pareto efficient paths are in general not Pareto efficient period-by-period.

13.5 The cosmopolitan solution

So far, we have assumed that countries behave non-cooperatively: although the citizens of each country consider the welfare of their descendants as morally relevant, they feel no obligation towards their contemporaries living in a different nation. This is probably a realistic assumption, at least to an extent, as shown for instance, in the context of environmental issues, by the difficulties in the promotion and application of the Kyoto Protocol. However, from an ethical perspective it is important to analyze the situation where equality of opportunity is at the centre of intergenerational *and* international issues, thus establishing a cosmopolitan normative benchmark.

In this section, we wish to find the *cosmopolitan lexicographic minimum solution*, $(\bar{I}, \bar{\tau})$, which solves the sequence of programmes $\mathbf{P}(n)$:

$$\max_{I_1^t, I_2^t, R^t, \tau^t} \min_{t \geq n} \left[u\left(R^t, \frac{(1-\tau^t)\alpha^t I_1^t}{f}\right), u\left(R^t, \frac{\beta^t I_2^t + \pi(\tau^t)\tau^t\alpha^t I_1^t}{1-f}\right)\right],$$

$\mathbf{P}_{\mathrm{C}}(n)$ subject to: $R^t = \rho R^{t-1} - I_1^t - I_2^t$, all $t \geq n$,

$I_1^t, I_2^t, R^t \geq 0$, all $t \geq n, 0 \leq \tau^t \leq 1$, all $t \geq n$, and R^{n-1} given.

where $\mathbf{P}_C(1)$ denotes the *cosmopolitan maximin programme* (MP).

Definition 4. A path $(\bar{I}, \bar{\tau})$ is a *cosmopolitan leximin solution* (CLS) if it solves $\mathbf{P}_C(n)$, all $n \geq 1$.

We can now characterize the dynamic path of utilities at a CLS.

Proposition 4. *At a CLS, $u_1^t = u_2^t = u^t$ and $u^t \leq u^{t+1}$, all $t \geq 1$.*

Proof:

1. At a CLS min $\{u_1^t, u_2^t\} \leq$ min $\{u_1^{t+1}, u_2^{t+1}\}$, all $t \geq 1$. In fact, if min $\{u_1^t, u_2^t\} >$ min $\{u_1^{t+1}, u_2^{t+1}\}$, some t, then by slightly decreasing I_1^t and I_2^t, it is possible to increase R^t and all utilities beyond t, without lowering min $\{u_1^t, u_2^t\}$ below min $\{u_1^{t+1}, u_2^{t+1}\}$. Hence, the original path was not leximin.
2. By part 1, the value of $\mathbf{P}_C(t)$ is reached at t. Hence, suppose $u_1^t > u_2^t$: it is possible to decrease I_1^t and increase I_2^t by a small amount, with $\mathrm{d}I_1^t = -\mathrm{d}I_2^t$. This raises the value of $\mathbf{P}_C(t)$, and hence the original solution was not leximin. QED.

By Proposition 4, we can focus on country 1's harvest and treat country 2 parametrically: at a CLS, I_2^t will be such that $u_1^t = u_2^t$, all t, given τ^t and I_1^t. Thus,

$$I_2^t = \frac{(1-f)(1-\tau^t)\alpha^t I_1^t}{f\beta^t} - \frac{\pi(\tau^t)\tau^t\alpha^t I_1^t}{\beta^t}.$$

Therefore, at a CLS,

$$I_1^t + I_2^t = I_1^t\left\{1 + \lambda^t\left[\frac{(1-f)(1-\tau^t)}{f} - \pi(\tau^t)\tau^t\right]\right\} = I_1^t\delta^t, \tag{13.20}$$

where $I_2^t \geq 0$ if and only if $\delta^t \geq 1$; that is, if and only if $(1-f)(1-\tau^t) \geq f\pi(\tau^t)\tau^t$. Given $\pi(\tau^t) \leq 1$, $\tau^t \leq 1-f$ is sufficient for the latter inequality to hold. Given (13.20), at a CLS,

$$R^t = \rho R^{t-1} - \delta^t I_1^t. \tag{13.21}$$

The next result characterizes the dynamic path of the transfer rate at a CLS.

Proposition 5. *Let $\lambda\pi(1) < 1$. Let $\hat{\tau}$ be defined by $(1-f)(1-\hat{\tau}) = f\pi(\hat{\tau})\hat{\tau}$. Let $\tilde{\tau}^t$ be defined by $1/\lambda^t = (1-\tilde{\tau}^t)\tilde{\tau}^t\pi'(\tilde{\tau}^t) + \pi(\tilde{\tau}^t)$. If $\tau\pi''(\tau) < 2|\pi'(\tau)|$, all $\tau \in [0, 1]$, then at all t, at a CLS either*

1. *$\tilde{\tau}^t \geq \hat{\tau}$, in which case $\tau^{t*} = \hat{\tau}$ and $I_2^t = 0$, or*
2. *$\tilde{\tau}^t < \hat{\tau}$, in which case $\tau^{t*} = \tilde{\tau}^t$ and $I_2^t > 0$.*

Proof:

1. Given (13.21), it is possible to write the utility of agent t in country 1 as $u(R^t, (1-\tau^t)\alpha^t(\rho R^{t-1} - R^t)/\delta^t f)$. Let us drop time subscripts.
2. Compute that the derivative of $(1-\tau)/\delta$ with respect to τ has the same sign as $\lambda\,(1-\tau)\tau\pi'(\tau) + \lambda\pi(\tau) - 1$; for any $\lambda > 1$, the latter expression is positive at $\tau = 0$, while – by our premise – it is negative at $\tau = 1$. Moreover

$$\frac{d[(1-\tau)\tau\pi'(\tau) + \pi(\tau)]}{d\tau} = 2(1-\tau)\pi'(\tau) + \tau(1-\tau)\pi''(\tau),$$

which is negative if $\tau\pi''(\tau) < 2|\pi'(\tau)|$, so that there is a unique τ such that

$$1/\lambda = (1-\tau)\tau\pi'(\tau) + \pi(\tau). \tag{13.22}$$

3. Therefore, given part 2, and noting that $(1-\tau)/\delta$ is concave at $\tilde{\tau}$, it follows that if $\tilde{\tau} \leq \hat{\tau}$, then $\tilde{\tau}$ is indeed a maximum, while if $\tilde{\tau} > \hat{\tau}$ then $(1-\tau)/\delta$ is maximized at $\hat{\tau}$. Let $\tau^* = \min\,\{\hat{\tau}, \tilde{\tau}\}$.
4. Suppose now that at a CLS $\tau^t \neq \tau^{t*}$, some t. By Proposition 4, the value of $\mathbf{P}_C(t)$ is reached at t. Then by parts 2–4 u^t, and thus the value of $\mathbf{P}_C(t)$ can be increased by choosing τ^{t*} while leaving R^{t+j} unchanged, for all $j > 0$. QED.

In the case where $\lambda^t = \lambda$, all t, we note that $\tilde{\tau}^t = \tilde{\tau}$, all t, and we define

$$\tau^* = \min\,[\hat{\tau}, \tilde{\tau}] \quad \text{and} \quad \delta^* = 1 + \lambda\left[\frac{(1-f)(1-\tau^*)}{f} - \pi(\tau^*)\tau^*\right] \tag{13.23}$$

Thus, the CLS transfer rate corresponds to the one-period Pareto optimal level. Proposition 5 has several implications. First, if $\lambda^t = 1$, then (13.22) implies $\tau^t = 0$ and $\delta^t = 1/f$. Second, Equation (13.20) suggests that in the cosmopolitan solution, we think of the choice of I_1^t as governed by intergenerational considerations, and the value of τ^t as determining intra-generational, international considerations. Proposition 5 tells us that τ^t is determined by the instantaneous international degree of technical advantage. In this sense, Proposition 5 allows us to separate considerations of intergenerational from international justice in the cosmopolitan case. Moreover given that at a CLS, $\tau^t = \tau^{t*}$, all t, for the sake of notational simplicity, in what follows we focus only on $\overline{R}$. Finally, τ^{t*} is the value that generation t in country 1 would choose selfishly, given the egalitarian constraint. The last observation suggests the following definition.

Definition 5. Generation t in country 1 is *unconstrained on* $\overline{R}$ if I_1^t and τ^t are the individually optimal values for t, given R^{t-1}, subject to the egalitarian constraint. If it is not unconstrained on $\overline{R}$, then we say that generation t in country 1 is *constrained*.

In other words, we say that a generation in country 1 is unconstrained if the prescribed path at date t is just the outcome of selfish optimization, given the egalitarian requirement. We immediately have:

Proposition 6. *If $\overline{R}$ is a CLS and agent t in country 1 is constrained, then I_1^t is lower than the individually optimal value, given the egalitarian requirement.*

Proof. Suppose not. Then, I_1^t could be decreased to the individually optimal value, thus increasing R^l and u^{l+1} for all $l \geq t$, a contradiction. *QED.*

Given Proposition 6, the following result can be proved.

Proposition 7. *If $\overline{R}$ is a CLS and $u^t < u^{t+1}$, some t, then agent t in country 1 is unconstrained.*

Proof. Suppose not. Agent t could increase u^t by increasing I_1^t a small amount. This raises the value of $\mathbf{P}_\mathrm{C}(t)$, and hence the original solution was not leximin. QED.

We will say that CLS's in which agents are unconstrained at all dates comprise Nirvana: for these are *intergenerationally just paths where each generation may optimize selfishly*. In Nirvana, the needs of future generations place no constraint on earlier generations' use of the forest.

Lemmas 3 and 4 provide sufficient conditions for a path $\overline{R}$, $\overline{\tau}$ to be the CLS.

Lemma 3. *Let $\overline{R}, \overline{\tau}$ be such that $u_1^t = u_2^t = u$, $\tau^t = \tau^{t*}$, $R^t \leq R^{t-1}$, and I_1^t is positive but lower than the selfishly optimal value, all t, given the egalitarian constraint. Then $\overline{R}$, $\overline{\tau}$ is the CLS.*

Proof:

1. Suppose not. Then, there is another feasible $\overline{R}', \overline{\tau}'$ such that $u_1'^t \geq u_1^t$ and $u_2'^t \geq u_1^t$, all t, with strict inequalities some t. Let u_R and u_H denote, respectively, the derivative of u with respect to the first and the second argument. Let $u_R^t = u_R(R^t,\ \alpha^t(1-\tau^t)I_1^t/f)$ and $u_H^t = u_H(R^t, \alpha^t(1-\tau^t)I_1^t/f)$ denote their value at t in the proposed path for country 1.
2. First, by Propositions 4 and 5, $\overline{\tau}' = \overline{\tau}$ and $I_2'^t = (\delta^{*t}-1)I_1'^t$, so that we only need to consider the path of the natural resource and we can focus on country 1. Next, as in Lemma 1, by strict quasi-concavity, $u_R^t < (\alpha^t(1-\tau^{*t})/f\delta^{*t})u_H^t$, all t, while $u_1'^t \geq u_1^t$, all t, implies u_R^t $[R'^t - R^t] + (\alpha^t(1-\tau^{*t})/f)u_H^t\ [I_1'^t - I_1^t] \geq 0$, all t, with strict inequality unless $(R'^t, I_1'^t) = (R^t, I_1^t)$.
3. Let t^* be the first generation for which $(R'^t,\ I_1'^t) \neq (R^t,\ I_1^t)$. Without loss of generality, we may assume $t^* = 1$. Then $I_1'^1 - I_1^1 = c^1 > 0$. Let $I_1'^2 - I_1^2 = c^2$; then $R'^2 - R^2 = -\rho\delta^{*1}c^1 - \delta^{*2}c^2$. Since $u_1'^2 \geq u_1^2$, $-u_R^2[\rho\delta^{*1}c^1] + [(\alpha^2(1-\tau^{*2})/\delta^{*2}f)u_H^2 - u_R^2]\delta^{*2}c^2 \geq 0$, which implies $c^2 \geq 0$, using the

inequality established in step 2. Iterating over t, we have that $u_1^{\prime t} \geq u_1^t$, implies $-\ u_R^t\ [\rho\delta^{*t-1}c^{t-1} + (\rho)^2\delta^{*t-2}c^{t-2} + \ldots + (\rho)^{t-1}\delta^{*1}c^1] + [(\alpha^t(1-\tau^{*t})/f\delta^{*t})u_H^t - u_R^t]\delta^{*t}c^t \geq 0$, which implies that $c^t \geq 0$, all t.

4. The rest of the proof is as in part 4 of Lemma 1. QED.

Finally, we can prove the following result.

Lemma 4. *Let $\overline{R}$, $\overline{\tau}$ be a path such that, at all t, $u_1^t = u_2^t = u^t$, $u^t \leq u^{t+1}$, and each agent in country 1 is unconstrained. Then $\overline{R}$, $\overline{\tau}$ is the CLS.*

Proof. u^1 cannot possibly be greater, and so $\overline{R}$ solves $\mathbf{P}_{\mathrm{C}}(1)$. Hence R^1 is determined. By induction, the path is the leximin path. QED.

As in section 13.3, we put more structure on the utility function in order to derive more precise results. Under (A1), if country 1 optimizes selfishly subject to the egalitarian constraint, the following first order conditions can be derived.

$$I_1^t = \frac{(1-b)\rho R^{t-1}}{\delta^{*t}},\ \text{all } t, \tag{13.24}$$

$$R^t = b\rho R^{t-1},\ \text{all } t. \tag{13.25}$$

where τ^{t*} and δ^{*t} are defined as in (13.23). Thus, country 1's indirect utility function is

$$u_1^t = (b\rho R^{t-1})^b \left(\frac{(1-\tau^{*t})(1-b)\alpha^t \rho R^{t-1}}{\delta^{*t} f}\right)^{1-b},\ \text{all } t. \tag{13.26}$$

By (13.25)–(13.26)

$$\frac{u_1^{t+1}}{u_1^t} = b\rho\left(\frac{(1-\tau^{*t+1})\alpha^{t+1}\delta^{*t}}{(1-\tau^{*t})\alpha^t\delta^{*t+1}}\right)^{1-b},\ \text{all } t. \tag{13.27}$$

We can now prove the following Theorem.

Theorem 4. *Let τ^* and δ^* be defined as in (13.23). Let $\alpha^t/\beta^t = \lambda > 1$, $\alpha^{t+1}/\alpha^t = \gamma > 1$, all t. Under (A1):*
A. If $\rho \geq \rho_{\min}^{CLS} = \frac{\gamma^{b-1}}{b}$, the CLS is given recursively by $\tau^t = \tau^*$, all t, and

$$R^t = b\rho R^{t-1}, I_1^t = \frac{(1-b)\rho R^{t-1}}{\delta^*}, I_2^t = (\delta^* - 1)I_1^t,\ \textit{all } t,$$

and no generation in country 1 is constrained. Moreover, if $\rho = \rho_{\min}^{CLS}$, then $u^t = u$, all t, and R^t decreases to zero; while if $\rho > \rho_{\min}^{CLS}$, then $u^t < u^{t+1}$, all t, and R^t decreases to zero if $\rho < 1/b$, remains constant if $\rho = 1/b$, and increases to infinity if $\rho > 1/b$.

B. If $\rho < \rho_{\min}^{CLS}$*, then the CLS is given recursively by* $\tau^t = \tau^*$*, all t, and*

$$R^t = \gamma^{b-1} R^{t-1}, I_1^t = \frac{\rho\gamma^{1-b} - 1}{\gamma^{1-b}\delta^*} R^{t-1}, I_2^t = (\delta^* - 1) I_1^t, \text{ all } t,$$

every generation is constrained, $u^t = u$*, all t, and* R^t *decreases to zero.*

Proof:

Part A. The result follows by Lemma 4. By (13.27) $u^t = u^{t+1}$, all t, if and only if $b\rho\gamma^{1-b} = 1$, while if $b\rho\gamma^{1-b} > 1$, individual optimization implies $u^t < u^{t+1}$, all t. The results concerning the dynamics of R^t follow from $\gamma > 1$ and (13.25).
Part B. The solution in part B implies $u^t = u$, all t, and since $(\rho\gamma^{1-b} - 1)/\rho\gamma^{1-b} < (1-b)$, I_1^t is less than the individually optimal level, all t. Then the result follows by Lemma 3. Finally, since $\gamma^{b-1} < 1$, $\lim_{t\to\infty} R^{*t} = 0$. *Q.E.D.*

Theorem 4 confirms that, if the structural differences between the two countries remain constant over time, sustainability in the literal sense of a non-shrinking forest size occurs if and only if love of the forest, b, is sufficiently great.

13.6 Cooperative bargaining

In this section, we study cooperative bargaining between the two countries over $(\bar{I}, \bar{\tau})$. Let $\Omega_j((\bar{I}, \bar{\tau}))$ be the lowest level of welfare for country j at $(\bar{I}, \bar{\tau})$. Let a *continuation path* $(\bar{I}, \bar{\tau})^n$ be a path of the three variables for all $t \geq n$. Let $F(R^{n-1}) = \{(\bar{I}, \bar{\tau})^n | R^t = \rho\ R^{t-1} - I_1^t - I_2^t,$ and $I_1^t,\ I_2^t,\ R^t \geq 0,$ all $t \geq n,$ given $R^{n-1}\}$ be the set of continuation paths feasible from R^{n-1}. Then, a *Nash bargaining solution* solves

$$\max_{(\bar{I},\bar{\tau}) \in F(R^0)} (\Omega_1((\bar{I}, \bar{\tau})) - m_1^1)(\Omega_2((\bar{I}, \bar{\tau})) - m_2^1). \qquad (\mathbf{B}(1))$$

where m_j^1 is the maximin value for j at the NCLS. Next, we denote the lowest level of welfare for country j at $(\bar{I}, \bar{\tau})^n$ as $\Omega_j((\bar{I}, \bar{\tau})^n)$ and define the following sequence of bargaining programmes.

$$\max_{(\bar{I},\bar{\tau})^n \in F(R^{n-1})} (\Omega_1((\bar{I}, \bar{\tau})^n) - m_1^n)(\Omega_2((\bar{I}, \bar{\tau})^n) - m_2^n), \qquad (\mathbf{B}(n))$$

where m_1^n, m_2^n are the maximin values for the two countries if they revert to the NCLS at $t = n$, given R^{n-1}. We define the leximin Nash bargaining solution as follows.

Definition 6. A path $(\bar{I}, \bar{\tau})$ is a *leximin Nash bargaining solution* (LNBS), if it solves $\mathbf{B}(n)$, all $n \geq 1$.

However, we do not here attempt to solve **B**(n). Instead, we solve the sequence $\{\text{NB}_t\}_{t=0,\ldots}$ of one-period problems

$$\max_{I_1^t, I_2^t, \tau^t} (u(R^t, H_1^t) - v_1^t)(u(R^t, H_2^t) - v_2^t), \text{ all } t. \qquad (\text{NB}_t)$$

where v_1^t, v_2^t denote the utilities that the two countries can reach at t if they revert to the NCLS *from t onwards*.[10] The solution of $\{\text{NB}_t\}_{t=0,\ldots}$ is interesting for two reasons. First, if $(\bar{I}, \bar{\tau})$ solves $\{\text{NB}_t\}_{t=0,\ldots}$ then it is a plausible candidate solution for **B**(n), all n. As for the other leximin problems analyzed in this chapter, the intuition is that if $(\bar{I}, \bar{\tau})$ solves the sequence of 'selfish' problems $\{\text{NB}_t\}$ with $u(R^t, H_j^t) \leq u(R^t, H_j^t)$, all t, $j = 1, 2$, and $v_1^1 = m_1^1$, $v_2^1 = m_2^1$, then it solves **B**(1). And similarly, for **B**(n).

Second, let R_{NB}^t, R_N^t, and R_C^t denote the value of R^t at the solution of $\{\text{NB}_t\}$, at the NCLS, and at the CLS respectively; and likewise for the other variables. Given our choice of disagreement point, we shall argue that the solution of $\{\text{NB}_t\}$ is interesting *per se* and not only as a candidate LNBS. By Theorem 3, noting that cooperative bargaining is Pareto efficient period-by-period, Proposition 8 immediately follows.

Proposition 8. *Let u(R, H) be any quasi-concave, differentiable utility function: at the solution to* $\{NB_t\}$, $\tau_{NB}^t = \tau_C^t$, *all t. Moreover, under (A1), if* $\rho \geq \rho_{\min}^{CLS}$ *then* $R_{NB}^t = R_C^t$, *all t.*

Thus, 'selfish' one-period Nash bargaining yields a transfer rate and a value of the natural resource equal to the CLS values, at all t. Proposition 8 does not depend on v_1^t and v_2^t, but if the latter are not specified I_1^t, I_2^t (and thus u_1^t, u_2^t) cannot be determined. Under (A1), the set of all Nash bargaining allocations of housing in period t can be written as $I_1^t = x_{NB}^t I^t = x_{NB}^t (1-b)\rho R^{t-1}$ and $I_2^t = (1 - x_{NB}^t) I^t = (1 - x_{NB}^t)\ (1-b)\rho R^{t-1}$ where x_{NB}^t is a number between zero and one, which can be viewed as reflecting the relative bargaining powers of the two nations (which we have assumed to be equal) and the disagreement point. Under (A1), the indirect utilities of the two countries at the solution of $\{\text{NB}_t\}$ are:

$$u_1^t = b^b (1-b)^{1-b} \left(\frac{\alpha^t (1 - \tau_{NB}^t)}{f} \right)^{1-b} (x_{NB}^t)^{1-b} \rho R^{t-1}, \text{ all } t, \qquad (13.28)$$

$$u_2^t = b^b (1-b)^{1-b} \left(\frac{\beta^t}{1-f} \right)^{1-b} [1 - x_{NB}^t (1 - \pi(\tau_{NB}^t)\tau_{NB}^t \lambda^t)]^{1-b} \rho R^{t-1}, \text{ all } t. \qquad (13.29)$$

13.7 International and intergenerational justice and sustainability

In this section, we compare the various institutional settings in terms of international and intergenerational equality of opportunity, welfare, and preservation of the natural resources, under (A1). We suppose that the structural parameters (describing preferences, technology, and the natural rate of growth of the forest) are unaffected by the institutional setting. First of all, given the theoretical focus on the relation between international and intergenerational equity, environmental concerns and human development, we analyze unconstrained leximin solutions: only on unconstrained paths is human development possible and the natural resource may not be depleted. Theorem 5 ranks the unconstrained CLS and NCLS in terms of their environmental and welfare properties.

Theorem 5. *Let $\alpha^t/\beta^t = \lambda > 1$, $\alpha^{t+1}/\alpha^t = \gamma > 1$, all t. Under (A1):*

A. For all $0 \leq b \leq 1$, if the NCLS is unconstrained then the CLS is unconstrained.
B. For all $0 \leq b < 1$, if $\rho \geq \rho_{\min}^{NCLS}$, then $(R_C^t/R_C^{t-1}) > (R_N^t/R_N^{t-1})$ and $R_C^t > R_N^t$, all $t \geq 1$.
C. For all $0 \leq b < 1$, if $\rho \geq \rho_{\min}^{NCLS}$, then $(u1_N^t/u2_N^t) > (u1_C^t/u2_C^t) = 1$, all t. Furthermore, $(u_{jC}^{t+1}/u_{jC}^t) > (u_{jN}^{t+1}/u_{jN}^t)$, all t, $j = 1, 2$.

Proof:

A. The result follows noting that $\rho_{\min}^{NCLS} \geq \rho_{\min}^{CLS}$ with strict inequality if $b < 1$.
B. The result follows from part A, (13.5) and (13.25).
C. By part A and by (13.6)–(13.7), $u1_N^t/u2_N^t = [\lambda(1-f)/f]^{1-b} > 1$, while $u1_C^t = u2_C^t$. Finally, the last part of the statement follows from (13.8) and (13.27). QED.

Theorem 5 states that if unconstrained solutions are considered, the noncooperative scenario is worse both from the viewpoint of international inequalities and from the environmental perspective, since it implies a lower growth rate *and* level of the natural resource at all t. Indeed, unsurprisingly, by Theorems 1 and 4.(A), if $b < 1$ sustainability in the literal sense of a nonshrinking forest size, is more likely in the cooperative scenario. Moreover, the noncooperative setting leads to a lower *growth rate* of welfare in both countries.

However, from the justice viewpoint, these cannot be arguments for the CLS, as far as the rich country is concerned, *if* the only obligation of a country is towards its own citizens. In fact, the possibility of welfare levels at $t = 1$ being higher at the NCLS, implies that there may be a conflict between welfare growth and sustainability, on the one hand, and intergenerational equality of opportunity, on the other hand. Theorem 6 characterizes such conflict.

Theorem 6. *Let $0 \leq b < 1$. Let τ^* and δ^* be defined as in (13.23). Let $\alpha^t/\beta^t = \lambda > 1$, $\alpha^{t+1}/\alpha^t = \gamma > 1$, all t. Under (A1), if $\rho \geq \rho_{\min}^{NCLS}$, then:*

A. $u2_C^t > u2_N^t$, all t;

B. *for country* 1, *there is a* $t_1 \geq 1$ *such that* $u1^t_C > u1^t_N$, *all* $t \geq t_1$, *and the unconstrained NCLS maximin dominates the unconstrained CLS if and only if*

$$\left[\frac{1}{2-b}\right]^{\frac{1}{1-b}} > \frac{(1-\tau^*)}{\delta^*}.$$

Proof:

Part A. By Theorem 5, comparing (13.7) and (13.26), and noting that $R^0_N = R^0_C$, we derive that $u2^0_C > u2^0_N$ if and only if

$$\frac{\lambda(1-\tau^*)(1-f)}{f\delta^*} > \left[\frac{1}{2-b}\right]^{\frac{1}{1-b}}.$$

By Theorem 4, the left-hand side of the inequality only depends on λ and f. Instead, the right-hand side is strictly decreasing in b and is maximized at $b=0$. Therefore in order to prove that $u2^0_C > u2^0_N$, it suffices to prove that $\frac{\lambda(1-\tau^*)(1-f)}{f\delta^*} > \frac{1}{2}$, or $\frac{\lambda(1-\tau^*)(1-f)}{f} > 1 - \lambda\pi(\tau^*)\tau^*$. However, by (13.23) the latter inequality is always true.

Part B. The first part of the statement follows from Theorem 5, (13.8) and (13.27). The second part follows from (13.6) and (13.26), noting that $R^0_N = R^0_C$. *QED.*

In other words, welfare levels in both countries eventually become higher at the CLS than at the NCLS, but it is possible that at the early stages of development the NCLS yields higher welfare in the rich country. (The unconstrained CLS always maximin dominates the unconstrained NCLS for the poor country.) This will depend only on three parameters: in fact, the left hand side of the inequality in Theorem 6.(B) depends on b, while by (13.23) the right-hand sides only depend on λ and f.

As concerns constrained paths, Theorem 2 characterizes constrained NCLS's with $1-b+b\gamma^{1-b} < b\rho\gamma^{1-b} < (2-b)$, while by Theorem 4 a constrained CLS exists if $b\rho\gamma^{1-b} < 1$. Therefore, it is interesting to analyze the cases where $\rho^{CLS}_{\min} < \rho < \rho^{NCLS}_{\min}$. Indeed, Theorem 7 proves that an even starker conflict between environmental and justice issues arises if the NCLS is constrained while the CLS is not.

Theorem 7. *Let* $b\gamma^{1-b} < 1$. *Let* $0 \leq b < 1$. *Let* τ^* *and* δ^* *be defined as in (13.23). Let* $(1-b+b\gamma^{1-b}/b\gamma^{1-b}) \leq \rho < \rho^{NCLS}_{\min}$. *Under (A1),*

A. $R^t_C > R^t_N$ *and* $(R^t_C/R^{t-1}_C) > (R^t_N/R^{t-1}_N)$, *all* $t \geq 1$. *Furthermore, there is a* $t_1 \geq 1$ *such that* $u1^t_C > u1^t_N$ *and* $u2^t_C > u2^t_N$, *all* $t \geq t_1$;
B. *Let* x_1 *be defined as in Theorem 2. For a given* $x \in [x_1,\ 1-x_1]$, *the constrained NCLS maximin dominates the unconstrained CLS for country 1, if and only if* $x^{\frac{b(\rho\gamma^{1-b}-1)}{1-b}} > (\frac{1-\tau^*}{\delta^*})(b\rho\gamma^{1-b})^{\frac{1}{1-b}}$; *and for country 2, if and only if*

$(1-x)\frac{b(\rho\gamma^{1-b}-1)}{1-b} > (\frac{(1-\tau^*)\lambda(1-f)}{\delta^* f})(b\rho\gamma^{1-b})^{\frac{1}{1-b}}$. However, for at least one country, the unconstrained CLS maximin dominates the constrained NCLS.

Proof:

Part A. Straightforward, from Theorems 2 and 4.
Part B. 1. At a constrained NCLS, we have

$$u_1^t = (\rho\gamma^{1-b} - 1)^{1-b} \left(\frac{x\alpha^t}{f}\right)^{1-b} \frac{R^{t-1}}{\gamma^{1-b}}, \text{ all } t,$$

$$u_2^t = (\rho\gamma^{1-b} - 1)^{1-b} \left(\frac{(1-x)\beta^t}{1-f}\right)^{1-b} \frac{R^{t-1}}{\gamma^{1-b}}, \text{ all } t.$$

The first part of the statement follows by comparing the latter expressions to (13.26) at $t=1$, noting that $R_N^0 = R_C^0$.

2. Suppose, contrary to the statement, that $x\frac{b(\rho\gamma^{1-b}-1)}{1-b} \geq (\frac{1-\tau^*}{\delta^*})(b\rho\gamma^{1-b})^{\frac{1}{1-b}}$ and $(1-x)\frac{b(\rho\gamma^{1-b}-1)}{1-b} \geq (\frac{(1-\tau^*)\lambda(1-f)}{\delta^* f})(b\rho\gamma^{1-b})^{\frac{1}{1-b}}$. Summing the two inequalities we obtain $\frac{b(\rho\gamma^{1-b}-1)}{1-b}(b\rho\gamma^{1-b})^{-\frac{1}{1-b}} \geq \left(1 + \frac{\lambda(1-f)}{f}\right)\frac{(1-\tau^*)}{\delta^*}$. Note that the left-hand side is strictly decreasing in $b\rho\gamma^{1-b}$. Hence, if the inequality holds for some value of $b\rho\gamma^{1-b}$, then it holds for $b\rho\gamma^{1-b}$ arbitrarily close to one. If $b\rho\gamma^{1-b}=1$ then $1 \geq \left(1 + \frac{\lambda(1-f)}{f}\right)\frac{(1-\tau^*)}{\delta^*}$ or equivalently $1 \geq \lambda\pi(\tau^*)$ which is always false by (13.23). QED.[11]

Our next exercise, then, is to study whether the possibility of bargaining can alter the trade-off between environmental and justice concerns. Recall that u_{jN}^t denotes the utility of generation t in country j along the original NCLS path, while v_j^t is the utility of generation t in country j if bargaining breaks down and both countries revert to the NCLS path *from t onwards*.

Theorem 8. *Let $0 \leq b < 1$. Let $\alpha^t/\beta^t = \lambda > 1$, and $\alpha^{t+1}/\alpha^t = \gamma > 1$, all t. Let τ^{*t} and δ^{*t} be defined as in (13.23). Assume (A1). If $\rho > \rho_{\min}^{CLS}$, then*

A. $u_{jNB}^{t+1} > u_{jNB}^t$, all t, j.
B. $R_{NB}^t > R_N^t$, all $t \geq 1$, and $u_{jNB}^t \geq u_{jN}^t$, all t, $j=1, 2$, with strict inequality for $t > 1$.
C. If $x_{NB} > 1/\delta^*$ then $u_{1NB}^t > u_{1C}^t$ and $u_{2NB}^t < u_{2C}^t$, all t, and the opposite holds if $x_{NB} < 1/\delta^*$. If $x_{NB} = 1/\delta^*$, then $u_{jNB}^t = u_{jC}^t$, all t, j.

Proof:

Part A. Assuming the bargaining power of the two countries to be constant over time and given that $\alpha^t/\beta^t = \lambda$, all t, implies $\tau_{NB}^t = \tau_C^t = \tau$, all t, and it can be shown that $x_{NB}^t = x_{NB}$, all t. Then the result follows from (13.28)–(13.29).

Part B. First, by Proposition 8 and Theorems 1 and 2, we have $R^t_{NB} \geq R^t_N$, all t, with strict inequality for $t > 0$. Then, it follows that $v^t_j \geq u^t_{jN}$, all t, j, with strict inequalities for $t > 1$. Then the result follows from $u^t_{jNB} \geq v^t_j$, all t, j.
Part C. Immediate from (13.28)–(13.29) since $\tau^t_{NB} = \tau^t_C$ and $R^t_{NB} = R^t_C$, all t. QED.

Theorem 8 proves that if the countries engage in cooperative bargaining, welfare growth, environmental issues and intergenerational concerns can be reconciled in the sense that the NB path leximin dominates the NCLS and yields increasing welfare in both countries, while the CLS path of the natural resource is chosen. (Hence, as in the CLS, in the bargaining solution sustainability in the literal sense obtains if and only if $\rho \geq 1/b$.) However, Theorem 8(C) states that in general international equality of opportunity – as formulated in the cosmopolitan view – will not hold and international inequalities will persist.

13.8 Conclusions

This chapter confirms the main qualitative conclusions of Roemer (ch. 12 in this volume) on the relationship between intergenerational and environmental issues and it generalizes them to a two-country world. A higher rate of technical progress, a higher love of the forest, and a higher growth rate of the forest, make Nirvana (intergenerationally just path with increasing utilities) *and* a sustainable path of the forest more likely. However, if generations are constrained, a higher rate of technical progress leads to a faster depletion of the forest on the intergenerationally just path.

This is true in all the institutional frameworks considered. In fact, as noted in the introduction, there are several contemporary views in political philosophy about what constitutes international justice. Our study does not adjudicate this debate; however, it contributes to the discussion by highlighting the implications of different views on international justice. In particular, two main conflicts are shown: one between environmental and justice (based on an equality of opportunity view of justice) concerns, and the other between international and intergenerational justice. In fact, as compared to the NCLS, the cosmopolitan EOp path yields a more sustainable path of the natural resource and it eliminates international inequalities, but it may yield a lower welfare level for the worst-off generations which conflicts with an intergenerational EOp view. By implementing the efficient level of extraction of the natural resource, the bargaining outcome solves the trade-off between sustainability and intergenerational justice, but the potential conflict between international and intergenerational justice persists.

It is worth pointing out that the above results are robust with respect to various changes in the assumptions. For instance, although we have focussed on a model where differences in development (more specifically, in technical knowledge) between the two countries persist, similar conclusions are reached if the two countries gradually converge over time, so that $\alpha^t/\beta^t = \lambda^t$ with $\lambda^t \to 1$ as $t \to \infty$. Furthermore, although only two countries have been considered, similar results can be obtained in a model with N identical poor countries which behave non-cooperatively among themselves (our model can be interpreted as depicting two coalitions of countries).

However, our results suggest at least three lines for further research, especially on international justice. First, although we have not directly tackled the philosophical issue of the appropriate concept of international justice, a fully-fledged theory of equality of opportunity would require a proper treatment of this issue. Actually, one could argue that the main forms of injustice of our time are those in the international arena, rather than within country borders. Second, from a formal viewpoint, it would be interesting to analyze more asymmetric forms of interaction, where the North faces a much less developed South with a considerably lower economic and political power. In this case, some Stackelberg-type solution concept of the leximin bargaining between nations would be necessary.

Finally, our model provides a further counter-example to the results on the stationarity of the maximin social welfare function applied to the intertemporal framework (e.g., Arrow 1973; Dasgupta, 1974; for a discussion see Silvestre, 2002). In our model the presence of the natural resource (the state variable) in the agents' utility function leads to the possibility of growth in the economy along the maximin path, even if agents do not care about future generations. Our assumption seems quite reasonable in environmental issues (and not only) and thus it would be interesting to provide a general characterization of intergenerational maximin paths with increasing welfare.

Notes

1. World Development Indicators Database, *World Bank,* 2004.
2. In other words, nature has a *use* value. However, we may also assume that it has an *existence* value.
3. By (1) the sustainable rate of harvest is $I_1^t + I_2^t = (\rho - 1)\, R^{t-1}$, all $t \geq 1$.
4. No resources are consumed in technological innovation in our model. We could complicate the problem in a number of ways, and in particular we might endogenize technological progress. However, investment in research and innovations is unnecessary to expose the problem we wish to concentrate upon, namely the issue of *intra*- and *inter*generational justice with exhaustible natural resources. For a treatment of endogenous technical progress in a similar context, see section 13.5.
5. Houses depreciate fully at each generation, and so the only bequest one generation leaves to the next is the size of the forest, and its technological knowledge.

6. This assumption guarantees that the leximin solutions analyzed in the rest of the chapter must entail positive consumption of both goods at every date for at least one country.
7. In sections 13.3 and 13.5, first, we assume that the value of all maximin and leximin programmes is attained. Then, when we adopt a specific functional form, we prove that the supremum is attained.
8. The pair $(\bar{I}_1, \bar{I}_2)$ could actually denote sequences of *functions* $I_j^t : \Re_+ \rightarrow \Re_+$ mapping the level of the state at t, R^{t-1}, onto a value of the harvest for country j at t. Indeed, the NCLS strategies in Theorems 1 and 2 define a *feedback* equilibrium.
9. An alternative proof of Lemma 1 can be derived as in theorem 1, if we also assume that the marginal rate of substitution along the putative LS_j is smaller than unity at all t.
10. Alternatively, v_1^t and v_2^t may be defined as the utilities that the two countries would have reached along the original NCLS path. However, this choice seems less convincing since it does not capture the actual disagreement point of the two countries and it may raise issues of time consistency.
11. By setting $x = 1 - x_1$, we prove that for the first inequality in Theorem 7(B) to hold for some x, it must be $1 > [(1 - \tau^*)/\delta^*]^{1-b}\, b\rho\gamma^{1-b}$. Similarly, by setting $x = x_1$, we prove that for the second inequality in Theorem 7(B) to hold for some x it must be $1 > [(1 - \tau^*)\lambda(1 - f)/f\delta^*]^{1-b} b\rho\gamma^{1-b}$. Each inequality can hold if $b\rho\gamma^{1-b}$ is sufficiently close to one.

References

Arrow, K. (1973) 'Rawls's Principle of Just Saving', *Swedish Journal of Economics*, vol. 75, pp. 323–35.

Dasgupta, P.S. (1974) 'On Some Alternative Criteria for Justice between Generations', *Journal of Public Economics*, vol. 3, pp. 405–23.

Nagel, T. (delivered 2004) 'The Problem of Global Justice', Storrs Lectures, Yale University Law School.

Pogge, T.W. (2002) *World Poverty and Human Rights: Cosmopolitan Responsibilities and Reforms* (Cambridge UK: Polity Press).

Rawls, J. (1999) *The Law of Peoples* (Cambridge, MA: Harvard University Press)

Silvestre, J. (2002) 'Progress and Conservation under Rawls's Maximin Principle', *Social Choice and Welfare*, vol. 19, pp. 1–27.

World Bank (2004) *World Development Indicators Database*.

14
Intergenerational Equity and Human Development

Joaquim Silvestre

14.1 Introduction

14.1.1 Intergenerational and intragenerational welfare

The last two hundred years have shown spectacular progress in the standard of living or quality of human life, at least for most people in the developed world, as evidenced by:

- the lengthening of the average lifespan, which has followed a linear trend in the last 160 years[1]
- the increase in the amount of education per person,
- the improvement in civil rights and the ability to exercise personal liberties,
- the growth in labour productivity, allowing a higher level of material consumption for the same input of labour time.

One can question whether this is also true of most people in developing countries. Indeed, the blatant intragenerational inequalities are the main qualification to any statement about human development, particularly across countries and in light of demographic expansions. Are the poorest of the poor better off today than fifty years ago? I honestly do not know. But, despite the importance of contemporaneous inequality, and the obvious connections between the two, here I focus on intergenerational, rather than intragenerational, welfare. Hence, for the sake of the argument, I abstract from intragenerational inequalities and demographic issues, and postulate a representative person in each generation.

It can be projected that the improvement in human welfare, often fuelled by the creation and implementation of technological knowledge, may continue in the future. But there are dissenting voices: the continued abuse of natural sinks beyond their regenerative capacities (climate change, oceans) may result in catastrophic failures of natural systems with serious effects on the quality of life. Yet I leave its analysis to more authorized voices, and focus on a benign view of future possibilities. 'Possibilities' is the key word: my discussion is normative: its underlying conjecture, or hope, is that some policies that satisfy sensible normative criteria entail the continuation of human progress. Of course, these policies may well require the reduction of emissions into crucial sinks or the conservation of natural environments.

Here I explore two criteria of intergenerational equity, namely maximin (leximin) and sufficientarianism, and the extent to which they are compatible with human development. Both criteria are anchored in the atemporal case, and their extension to the intergenerational context raises difficult issues.[2] Based on the idea of minimizing the effects on welfare of circumstances beyond a person's control, and given the moral irrelevance of the date of birth, both the Rawlsian maximin (or leximin) principle and the less demanding sufficientarianism criterion (which requires 'good enough' standards for each generation) advocate policy measures that benefit the early generations, since they likely are the worst off in a world with technological progress. But transferring resources to them may jeopardize future standards of living. Sections 14.3 and 14.4 discuss the implications of these criteria on human development.

The analytical method followed here has two main features.

(a) The intertemporal economy is modelled as an overlapping-generations configuration with relatively long periods. The length of a period is interpreted as that of a life stage (such as youth, maturity or seniority). Generations in their nonproductive life stages coincide with those in their productive stages, so that in principle both backward transfers (from the currently productive to the no longer productive) and forward transfers (from the currently productive to the not yet productive) are feasible. For simplicity, however, the main text just contemplates the coexistence of two generations – a mature and a senior one – so that only backward transfers are feasible, whereas, on the contrary, the appendix covers a case of only forward transfers.

(b) Economic decisions involving the present and the future have to be made at a given period. The past cannot be changed, but at the decision period an old generation is present whose standard of living cannot be totally controlled.

14.2 Leximin and sufficientarianism as intergenerational criteria

14.2.1 The conflict between leximin and human development

For maximin and leximin, the basic intuition is that the dependence of a person's quality of life on circumstances beyond her control should be reduced as much as possible. Therefore, society should minimize, within the realm of possibility, the effects of bad luck on the quality of life, where 'bad luck' is understood in a relative sense, i.e., compared to that of other members of society. This leads to maximizing the allocation of 'primary goods' to the people with the least of them. Modulo incentives, maximin tends to imply equality (say, of primary goods, or of opportunity) in an atemporal society.

One difficulty with the intergenerational maximin is that it tends to impose the equality of the standards of living across generations. The logic parallels that of the atemporal case: if there is human development, then later generations are better off than earlier ones. Because the date of birth is surely beyond a person's control, neutralizing the effects of bad luck, where 'bad' is understood relative to the luck of other generations, the welfare of present generations should be increased up to the point where future generations are no better off than the present generation, so negating any improvements in the standards of living at a leximin solution. Thus, a basic tension appears between the intuition for leximin and the natural desideratum of human development.

14.2.2 Dynasties vs human development as a public good

The conventional solution to the dilemma appeals to altruistic utility functions that include as arguments the utility, welfare or consumption of future generations, e. g., instead of maximizing the minimum standard of living, one maximizes the minimum of a utility function where the utility of future generations is itself an argument in the current generation's utility. Maximizing the minimum of utility so understood may well imply an increasing path of standards of living.

What are the empirical bases for this altruism towards future generations? One view is that it reflects the concern for one's own descendants. Many parents are happy to sacrifice themselves for the sake of their children's welfare: this leads to a 'dynastic' view of utility which is actually close to John Rawls's own treatment of intergenerational welfare (1971, Section 44). But, at least if taken literally, the dynastic utility would not apply to childless individuals, who would then become the worst off. Should we then transfer income from parents to childless people?

A more universal justification for altruism towards future generations appeals to the notion of *human development as a public good*: we may feel justifiably proud of mankind's recent gains in, say extraterrestrial travel, or average life expectancy, and wish them to continue into the far future even at a personal cost. Indeed, there seems to be an asymmetry in the way we feel about contemporaneous vs temporally disjoint inequality: a person in a poor country may be reluctant to sacrifice her standard of living for the sake of improving that of a person in a *richer country*, while at the same time she may be willing to make some sacrifices for the welfare of unrelated, yet-to-be born people who will in all likelihood be richer than her.

14.2.3 The standard of living function in leximin

Even if people do derive subjective utility from the welfare of future generations, the question remains of which is the normatively relevant maximand in the maximin (or leximin) calculus. It can be argued that it should be the standard of living or quality of life of each person, without reference to the subjective utility that she may derive from the welfare of other people. This seems to be Rawls's position when justifying the (contemporaneous) 'difference principle'. Accordingly, here I focus on the maximin (leximin) principle as applied to an index of a person's own welfare, quality of life or standard of living, in the spirit of Amartya Sen's 'capability', or G. A. Cohen's 'midfare.'[3] More operationally, we can consider the Human Development Index (HDI) produced by the United Nations Development Program, which considers three dimensions: (a) long and healthy life, (b) education, and (c) consumption (or GDP per capita), to which (d) leisure time and (e) environmental quality could be added.

In order to avoid confusion, I refrain in what follows from referring to such an index as 'utility' or even 'welfare,' and I settle on the term *standard of living*, to be denoted by the Greek letter Λ instead of the customary U or W, which would suggest 'utility' or 'welfare.'

We understand the standard of living of a generation as an aggregate index of the (opportunities to access) various primary goods such as available to people born in a certain time interval (say, the early 1990s) at the different periods of their life. Note that such a numerical index aggregates over two dimensions:

(a) over the different goods, such as health, education, or environmental amenities, in addition to consumption, and
(b) over the different stages of life, such as childhood, early adulthood, maturity and seniority.

The two dimensions are not totally independent, since education is concentrated in the early stages, and health services in the latter ones.

Roemer (1996, Section 4.2) provides a clarifying discussion of the issues raised by the aggregation over Rawlsian primary goods, but the aggregation over life stages presents its own challenges: see Dennis McKerlie (1989) and Larry Temkin (1993).[4] A major theme in the present chapter is the extreme lifetime pattern that leximin may impose due to aggregation over life stages (see in particular section 14.5.1 below).

14.2.4 Factors attenuating the leximin–human development conflict

The conflict between leximin and human development may not be as serious as it looks at first sight. The actual degree of compatibility between the two desiderata is ultimately an empirical question, having to do with the real-world parameters of the various structural functions (for technological change, and investment, positive and negative, in physical, human and natural capital, and socially accepted minimal or safety levels). Several factors may contribute to reduce the tension between maximin and human development.

One such factor is the presence of intergenerational public goods, such as stocks of knowledge or natural environments. Because human development operates to a large extent through the creation of knowledge, an intergenerational public good, the conflict between present consumption and future productive capacity is not as hard as traditional growth models would suggest: In particular, the act of providing for the consumption needs of the older generations entails the investment in forms of capital than enhance both the present and the future productivity of labour, creating a positive intergenerational externality. As illustrated in section 14.4 below, if the economy is sufficiently productive, and the initial level of capital is sufficiently high, then leximin implies human development. Other types of intergenerational public goods have been analyzed in Silvestre (2002) and in Roemer (2007).

A second factor is the presence of incentive-based upper limits on the amount of resources that can be transferred backwards. As highlighted in the current debates on the 'sustainability' of social security systems in Japan, the United States and Europe, backward transfers (including investment and consumption programmes that benefit the elderly) usually require the taxation of the earlier generations. Incentive constraints may impose ceilings to the amounts that can be raised by taxing the income of the currently productive workers, limiting the feasible extent of backward transfers.

In addition to these factors, which are based in either physical or behavioural realities, socially-determined bounds on various transfers and investments may also exist which mitigate the conflict between leximin and human development when included in the list of constraints in the leximin calculus: see section 14.5.1 below. Under these additional constraints, and if the economy is sufficiently productive, the leximin solution may yield increasing paths of standards of living (see Silvestre, 2002, as well as Roemer and Roberto Veneziani, 2004, described in the Appendix).

14.2.5 The everything-to-seniors problem

If the leximin solution does entail human development, then the lifetime consumption pattern may display a strong asymmetry among the various stages of life: the productive efforts of those in the working ages are devoted to investing and to satisfying the consumption needs of the older generations, delaying their own consumption until their seniority period. With either exogenous or endogenous growth, the leximin criterion prioritizes the needs of the currently senior. Barring satiation in the standard-of-living function relative to consumption at the late phases of life, or strict complementarities among consumption at the different life stages, it favours:

- large backwards transfers, such as large pay-as-you-go social security programmes, and large government debt.
- large investment in medical research and infrastructure, and the provision of health services to the aging.
- low forward transfers, and, in particular, low supply of education,
- to the extent that there are investment–consumption tradeoffs, low investment in physical, human and natural capital.

This extreme pattern becomes particularly distressing if there is uncertainty about the length of life, so that a fraction of the population dies prematurely: Section 14.5.1 below comments.

14.2.6 Sufficientarianism

Sufficientarianism (Harry Frankfurt, 1987, see also Elizabeth Anderson, 1999, Richard Arneson, 2000, 2002, 2005, Martha Nussbaum 2000, and Roemer, 2004) weakens the egalitarian implications of maximin by considering only *absolute*, rather than relative, bad luck, i.e., it advocates the minimization, within the realm of possibility, of the negative effects of absolute bad luck on the quality of life. It implies guaranteeing everybody *good-enough*, or *sufficient* levels of primary goods, but it allows for some people to have a lot more than what is sufficient. A programme ensuring that nobody falls below the poverty level can be viewed as an application of sufficientarianism, but for a low sufficiency floor. In principle, sufficientarianism may aim at guaranteeing relatively high levels of primary goods to everybody.

Mathematically, the leximin criterion selects the solution path (or paths) to a (sequential) constrained optimization problem, where each term of a sequence of objective functions is maximized over the corresponding constraint set, iteratively defined by all relevant physical, resource and possibly incentive restrictions. Sufficientarianism, on the contrary, defines a set of admissible paths by a list of inequality constraints involving not only these physical and incentive-based restrictions, but also ethically-motivated lower bounds (the 'good enough' levels) on various primary goods or other functions of the paths. In this way sufficientarianism formally appears as a relaxation of leximin to the extent that leximin solutions satisfy 'good enough' restrictions and, hence, they belong to the sufficientarianist set.

So defined sufficientarianism tends to be indeterminate, because many paths typically satisfy its admissibility conditions. In practical terms, however, sufficientarianism is often understood as a weakening of the redistribution directives of leximin, requiring less public intervention, and being implemented by a minimal deviation from *laissez-faire*.

Contrary to the leximin approach, sufficientarianism faces the challenge, both in the atemporal and the intergenerational applications, of specifying 'good-enough' levels in a nonarbitrary manner (see Arneson, 2002). In the atemporal context, the 'good-enough' level can be defined either in absolute or in distributive terms: relative sufficiency in some good or index of goods is most naturally defined in terms of its distribution in the population, e. g., as a percentage of its median or mean, or in Arneson's (2002, p. 173) words, 'it might be stipulated that everyone has enough income and wealth when nobody has less than some fraction of the average level.' For instance, we may view as a good-enough life expectancy 70 per cent of the current one in Japan.

Focusing on the distribution of a good or index of goods among the various generations that coexist in a given period poses several problems. First, the weight of many primary goods changes across the lifetime: as just observed, education is concentrated on the early years, and medical care in the later ones. Second, this approach leaves too many degrees of freedom, and some solutions are not too interesting. For instance, any constant path, even at very low levels, automatically satisfies this form of sufficientarianism.

In an intertemporal context even 'absolute' good-enough levels should refer to the current technological and resource possibilities. What could have been a sufficient level of health care in 1930 is no longer sufficient in 2005. If, for example, health technology improves so that the life expectancy of the population as a whole increases, then the 'sufficient' level of life expectancy for each subgroup should also increase. Section 14.3 below applies this notion to a simple OLG with exogenous technological growth. Some of the ideas developed there are extended to endogenous technological growth in section 14.5.2 below.

Independently of any intuitive attractiveness, or lack of it, sufficientarianism offers in the intergenerational context some practical advantages over leximin. First, it is more development-friendly. Second, it allows for more symmetric lifetime paths, less biased in favour of the late stages of life, than those of leximin when it is compatible with human development. Third, because it is formulated as a list of constraints rather than as an optimization problem, it can be applied in a disaggregated manner primary good by primary good, as in Michael Walzer (1983), or Nussbaum (2000). Furthermore, the floors can be disaggregated life stage by life stage, and state of the world by state of the world, taking into account the resolution of any uncertainty as it appears. We find an application of this advantage in section 14.3, when comparing leximin with sufficientarianism in an overlapping generations model: because at the planning date an old generation (Generation Zero) exists, the welfare of which involves past variables that can no longer be controlled, leximin requires as an initial condition the current obligations to that generation, whereas under sufficientarianism we can apply the same age-specific sufficiency norms (perhaps date dependent, to account for enhanced technological possibilities) to all generations, including Generation Zero.[5]

The next section analyses, using a simple, exogenous growth model, the implications of alternative sufficiency criteria that take into account the increasing technological possibilities. In particular, there is a best sufficiency level in the simple model considered. Moreover, it compares the leximin and sufficientarianism paths of standards of living taking account of the initial conditions, which are a basic datum for meaningful comparisons.

14.3 Leximin vs sufficientarianism: initial obligations and the 'everything-to-seniors' problem

14.3.1 An OLG model with exogenous growth

To simplify matters, not only do I postulate a single representative consumer in each generation, and abstract from uncertainty, as before, but moreover, and contrary to the previous discussions, I:

(a) consider a single good, which could be interpreted as an index of primary goods, *à la* Human Development Index, say an index of health services, consumption, knowledge ...
(b) abstract from any bounds on backwards transfers other than nonnegativity; In particular, because I do not impose safe minimum standards on education or the environment, I consider an OLG model without young people to be educated;
(c) model technological progress as exogenous growth: the (composite) good is produced by a fixed labour supply that becomes increasingly productive.

Thus, I posit an OLG model where time goes by in relatively long periods, numbered $t = 1, 2 \ldots$ and is populated by generations, also numbered $t = 1, 2, \ldots$. For $t = 1, 2, \ldots$, Generation t lives over two periods, being *mature* in period t and *senior* in period $t + 1$; Its consumption vector is denoted (c_{mt}, c_{st}), with associated standard of living $\Lambda(c_{mt}, c_{st})$.

The function Λ plays a central role in normative analysis. As discussed in section 14.2 above, it should be understood as a lifetime standard of living, although there are conceptual difficulties if the date of death is uncertain. We could consider an alternative definition, as period-specific standards of living, in the spirit of the sufficientarianist 'good-enough' levels, but leximin is ill suited for that, because it is based on the optimization of an objective function.

For simplicity, I formalize the production period as having the same time span as the lifetime periods. We can visualize that people at maturity work a fixed amount of time, normalized to $L_t = 1$, and the productivity of their labour increases at a constant rate. Let the amount of the good available in period t be β^t, $\beta > 1$.

I posit an overlapping generations (OLG) economy with a senior generation to which the currently productive workers can transfer produced goods. In addition, I envisage that a decision involving the present and the future has to be made at a given point in time, but without the ability to change the past. Some of the issues studied here cannot be considered in non-overlapping-generations models (such as Roemer, 2007), or models without 'senior' generations (i.e., models where forward transfers are perhaps possible, but backward transfers are not, such as Roemer and Veneziani, 2004) or overlapping–generations models without a definite decision period (such as Tomoichi Shinotsuka et al., 2007). Accordingly, consider an overlapping-generations model with the temporal structure of Figure 14.3.1, starting in Period One, which is the decision period, where a senior generation (Zero) and a mature generation (One) coexist.

14.3.2 Suficientarianism relative to the production possibilities

In period t, the amount of output is β^t and a senior and a mature generation coexist. Let us postulate that the sufficient level of the good in period t for the currently senior (resp. mature) generation is defined as a fraction σ (resp. μ) of current output, where $\sigma + \mu \leq 1$. Because it is the currently mature generation who produces the output, any part of it allocated to the senior generation is transferred backwards. Assume that lump-sum, costless backward transfers are feasible, or, using Arthur Okun's expression, that the transfer buckets do not leak. Suppose that, in the spirit of the minimal deviation from *laissez-faire* alluded to in section 14.2.6 above, the planner wishes to minimize backward transfers subject to sufficiency. She then sets $c_{s,t-1} = \sigma\beta^t$ and $c_{mt} = (1-\sigma)\beta^t$, $t = 1, 2, \ldots$ Therefore, the standard of living of Generation t, $t = 1, 2, \ldots$ is $\Lambda((1-\sigma)\beta^t,\ \sigma\beta^{t+1})$. Note that $c_{s0} = \sigma\beta$, i. e., the sufficiency criterion is also applied to the senior-age consumption of Generation Zero, even though its standard of living is left undefined. The whole path of standards of living is governed by the fraction σ of output to be allocated to seniors that society considers sufficient. Clearly, the path is increasing in t.

The standard of living of Generation One, $\Lambda((1-\sigma)\ \beta,\ \sigma\beta^2)$, is a function of σ. Consider two societies identical in all respects except for their values of σ, which are $\overline{\sigma}$ and $\overline{\overline{\sigma}}$. If $\Lambda((1-\overline{\sigma})\beta,\ \overline{\sigma}\beta^2) > \Lambda((1-\overline{\overline{\sigma}})\beta,\ \overline{\overline{\sigma}}\beta^2)$, and Λ is homothetic, then the whole path for standards of living is higher in the society with $\overline{\sigma}$ than in that with $\overline{\overline{\sigma}}$: indeed, homotheticity implies that there is a strictly increasing function f such that the composition of f and Λ is homogeneous of degree r, hence

$$\begin{aligned}
&\Lambda((1-\overline{\sigma})\beta^t, \overline{\sigma}\beta^{t+1}) > \Lambda((1-\overline{\overline{\sigma}})\beta^t, \overline{\overline{\sigma}}\beta^{t+1})\\
&\quad\Leftrightarrow f(\Lambda((1-\overline{\sigma})\beta^t, \overline{\sigma}\beta^{t+1})) > f(\Lambda((1-\overline{\overline{\sigma}})\beta^t, \overline{\overline{\sigma}}\beta^{t+1}))\\
&\quad\Leftrightarrow \beta^{(t-1)^r} f(\Lambda((1-\overline{\sigma})\beta, \overline{\sigma}\beta^2)) > \beta^{(t-1)r} f(\Lambda((1-\overline{\overline{\sigma}})\beta, \overline{\overline{\sigma}}\beta^2))\\
&\quad\Leftrightarrow \Lambda((1-\overline{\sigma})\beta, \overline{\sigma}\beta^2) > \Lambda((1-\overline{\overline{\sigma}})\beta, \overline{\overline{\sigma}}\beta^2).
\end{aligned}$$

	Period 1 Output $= \beta$	*Period 2* Output $= \beta^2$	*Period 3* Output $= \beta^3$	*Period 4* Output $= \beta^4$	...	*Period t* Output $= \beta^t$	*Period t + 1* Output $= \beta^{t+1}$	...
Gen. 0	Senior c_{s0}							
Gen. 1	Mature $c_{m1} = \beta - c_{s0}$	Senior c_{s1}						
Gen. 2		Mature $c_{m2} = \beta^2 - c_{s1}$	Senior c_{s2}					
Gen. 3			Mature $c_{m3} = \beta^3 - c_{s2}$	Senior c_{s3}				
...				...	...	...		
Gen. t						Mature $c_{mt} = \beta^t - c_{s,t-1}$	Senior c_{st}	
...							...	...

Figure 14.3.1 Overlapping generations with exogenous growth

Moreover, if $\overline{\sigma} > \overline{\overline{\sigma}}$ (resp. $\overline{\sigma} < \overline{\overline{\sigma}}$) Generation Zero gets higher (resp. lower) consumption of the good in the society with $\overline{\sigma}$ than in that with $\overline{\overline{\sigma}}$.

If Λ is actually homogeneous of degree r, then $\Lambda((1-\sigma)\beta^t,\ \sigma\beta^{t+1}) = \beta^{tr}$ $\Lambda((1 - \sigma), \sigma)\beta)$, i. e., while output grows at rate β, the standard of living grows at rate β^r, independently of σ (or μ).

This leads to the question: which society would be best in the sense of generating the highest path of standards of living for Generations 1, 2, ...? Because a higher σ implies that more period-one output goes to Generation Zero, and thus that less is available Generation One, it might be mistakenly thought that σ should be low. In fact, under homotheticity the question reduces to choosing $\sigma \in [0, 1]$ in order to maximize $\Lambda((1-\sigma)\beta, \sigma\beta^2)$.

In the conventional case where $\Lambda(c_m,\ c_s) = u(c_m) + \theta u(c_s)$, with $\theta \leq 1$ and $u'' < 0$, the first-order equality is $u'((1-\sigma)\beta) = \theta u'(\sigma\beta^2)\beta$, i.e., $\frac{u'((1-\sigma)\beta)}{u'(\sigma\beta^2)} = \theta\beta$. If $\theta\beta > 1$, then $(1-\sigma)\beta < \sigma\beta^2$, i.e., $\sigma > \frac{1}{1+\beta}$ which is slightly below $\frac{1}{2}$, and σ itself may well be higher than ½, depending on u.

As an illustration, let $\beta = 1.1$ and $\Lambda(c_m, c_s) = \sqrt{c_m} + \sqrt{c_s}$. Then the σ that maximizes the standard-of-living path is $\sigma^* \equiv \frac{\beta}{1+\beta} = 0.5238$. Let us refer to sufficientarianism with parameter σ^* as the 'best sufficientarianism.'

Under best sufficientarianism, the consumption vector of Generation t is $((1-\sigma^*)\beta^t, \sigma^*\beta^{t+1}) = \beta^t((1-\sigma^*), \sigma^*\beta) = \beta^t(0.4762, 0.5762)$, with consumption at seniority quite larger than consumption at maturity, leading to the standard of living $1.4491\sqrt{1.1^t}$. If, on the contrary, σ is lowered to $\frac{1}{2}$, then the consumption vector of generation t is $\beta^t(0.5, 0.55)$ with consumption at seniority exceeding that of maturity exactly in line with productivity growth, but with lower standard of living $(\sqrt{0.5} + \sqrt{0.55})\sqrt{1.1^t} = 1.4488048\sqrt{1.1^t}$, for all generations, as well as lower backward transfer to Generation Zero. See the 'best sufficientarianist' curves in Figures 14.3.2 and 14.3.3.

14.3.3 The role of the initial obligation to seniors in leximin

The objective of the leximin planner is to choose a path that, first, maximizes the standard of living of the worst-off generation, then, among the solutions to that problem, maximizes the standard of living of the second-worst-off generation, and so on: in short, to leximin the standards of living. Because Generation Zero has never been mature, we must disregard its standard of living and, instead, assume that Generation Zero is exogenously entitled to a given amount c_{s0} of the good produced in period one, so that c_{s0} is now the initial condition of the model. We write a path in this economy as $(c_{s0}$, $\{(c_{mt},\ c_{st}),\ t = 1, 2, \ldots\})$. Note that $c_{m1} = \beta - c_{s0}$ is determined by the initial obligations, and hence the first policy variable is $c_{s1} \in [0,\ \beta^2]$, a period-2 variable, chosen by the planner in order to leximin the standards of living subject to $c_{m1} = \beta - c_{s0}$. The chosen path gives Generation One the standard of living $\Lambda(\beta - c_{s0},\ c_{s1})$.

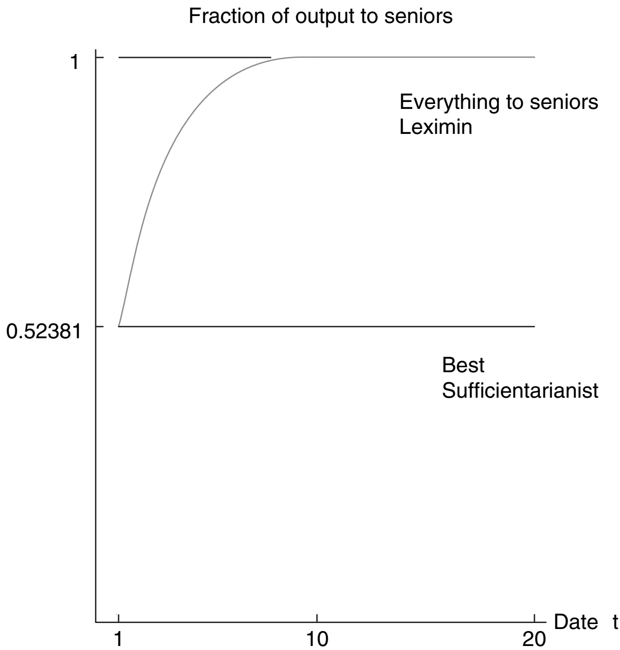

Figure 14.3.2 Fraction of output for seniors' consumption in the various solutions

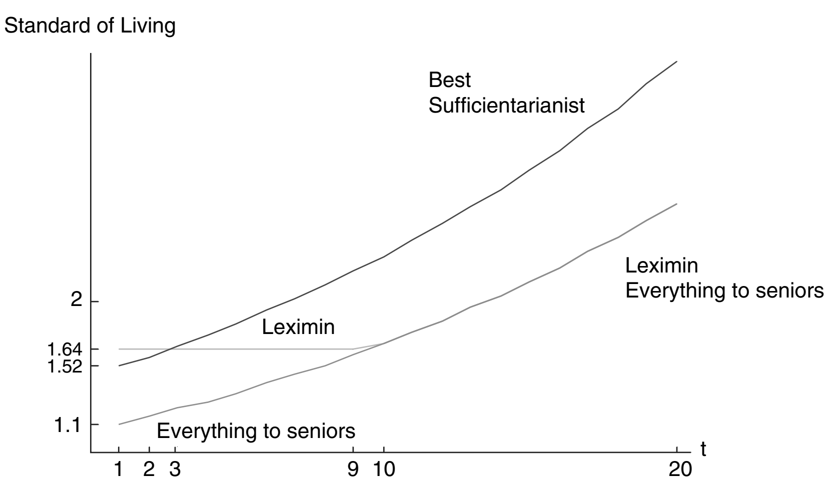

Figure 14.3.3 Paths of standards of living in the various solutions

Contrary to what was argued in section 14.2 above, suppose that there are no limits to backward transfers, so that in principle all the output produced in a given period can be costlessly transferred to the current seniors. Obviously, the leximin solution will depend on the level c_{s0} of the initial obligations to seniors.

As an extreme case, assume that senior Generation Zero is entitled to the whole amount of output produced in period one, i. e., namely $c_{s0}=\beta$. The leximin solution is then the path $(\beta,\ \{(0,\beta^{t+1}),\ t=1,2,\ldots\})$, to be called the 'everything-to-seniors' path: see Figure 14.3.3 for an illustration in the example of section 14.3.3 above. It imposes a relatively low standard of living, $\Lambda(0,\beta^2)$, for Generation One, and yields a lifetime consumption profile where consumption at maturity is constant at a very low (zero) level, where consumption at seniority grows in pace with technological progress, and, in the case where Λ is homogeneous of degree r, the standard of living grows at the rate β^r-1, i.e., there each generation is strictly better off than the previous one, and therefore, there is strict human development in the whole path, but it is a path that starts at a relatively low standard of living for Generation One.

Of course, this path is equivalent to the sufficientarianist path with $\sigma=1$, and, consequently, $\mu=0$, but it would not satisfy the sufficientarianist constraints were $\mu>0$.

As an initial obligation to seniors, $c_{s0}=\beta$ is unnaturally extreme, as would be a sufficiency standard requiring all output to be transferred to seniors. But if $c_{s0}<\beta$, then the feasible path $(\beta,\ \{(0,\ \beta^{t+1})?\ t=1,2,\ldots\})$ is no longer maximin.[6] For the rest of this section, assume that Λ is continuous, strictly increasing in c_s, nondecreasing in c_m, and, moreover, that there exists a real number c' such that $\Lambda(0,c')\geq\Lambda(\beta,\beta^2)$. In order to compute the leximin path when the initial obligation to seniors is $c_{s0}<\beta$, consider the following recurrence equation with two initial conditions, c_{s0} and c_{s1}, where c_{s0} is a datum (fixed in what follows) and c_{s1} will eventually be endogenously determined by solving the leximin programme.

Initial conditions: $c_{s0}\in[0,\beta)$, $c_{s1}\in[0,\beta^2]$. They define the standard of living of Generation One as $\Lambda(\beta-c_{s0},c_{s1})\equiv\Lambda_1$.

In order to write the laws of motion, define the function

$$q:\Re_+\to\Re_+:q(c_m)=\begin{cases}0,\ \text{if}\ \Lambda(c_m,0)\geq\Lambda_1,\\ \text{the unique } c_s \text{ that solves '}\Lambda(c_m,c_s)=\Lambda_1\text{,' if}\\ \Lambda(c_m,0)<\Lambda_1.\end{cases}$$

The above assumptions on Λ imply the following properties of q.

(I) q is well defined, because $\Lambda(c_m,c')\geq\Lambda(0,c')\geq\Lambda(\beta,\beta^2)\geq\Lambda(\beta-c_{s0},c_{s1})\equiv\Lambda_1$. Hence, if $\Lambda(c_m,0)<\Lambda_1$, then there exists a $c_s\in[0,c']$ that solves the equation '$\Lambda(c_m,c_s)=\Lambda_1$', and uniqueness follows from the fact that Λ is strictly increasing in c_s.

(II) q is bounded (by c').

(III) q is continuous, which can be argued as follows. Define $Q_0 \equiv \{(c_m, c_s) \in \Re^2_+ : \Lambda(c_m, c_s) \geq \Lambda_1\} \cap \{(c_m, c_s) \in \Re^2_+ : c_s = 0\}$, and $Q_1 \equiv \{(c_m, c_s) \in \Re^2_+ : \Lambda(c_m, c_s) = \Lambda_1\}$. The continuity of Λ implies that both Q_0 and Q_1 are closed, and thus so is $Q \equiv Q_0 \cup Q_1$. It is easy to check that Q is the graph of q (note that if $(c_m, c_s) \in Q_1$, then $\Lambda(c_m, c_s) = \Lambda_1$, which implies that $c_s = q(c_m)$ both in the obvious case where $\Lambda(c_m, 0) < \Lambda_1$, and in the case where $\Lambda(c_m, 0) \geq \Lambda_1$, because then the fact that $\Lambda(c_m, c_s) = \Lambda_1$ and that Λ is increasing in c_s implies that $c_s = 0$, i.e., $q(c_m) = 0 = c_s$). Thus, the graph of q is closed, which, together with the boundedness of q, implies that q is continuous.

(IV) q is nonincreasing, because if $\bar{c}_m < \bar{\bar{c}}_m$ and $q(\bar{c}_m) < q(\bar{\bar{c}}_m)$, then $0 < q(\bar{\bar{c}}_m)$, which implies that $\Lambda_1 = \Lambda(\bar{\bar{c}}_m, q(\bar{\bar{c}}_m)) > \Lambda(\bar{\bar{c}}_m, q(\bar{c}_m)) \geq \Lambda(\bar{c}_m, q(\bar{c}_m))$, contradicting the fact that $\Lambda(c_m, q(c_m)) \geq \Lambda_1, \forall c_m$, by the definition of q.

Laws of motion:

Note that, by assumption, $\beta - c_{s0} \geq 0$ and $\beta^2 - c_{s1} \geq 0$. We write $\hat{c}_{s1}(c_{s1}) = c_{s1}, \forall c_{s1} \in [0, \beta^2]$ (i.e., $\hat{c}_{s1}$ is the identity function.) Define

$$\hat{c}_{s2}(c_{s1}) \equiv q(\beta^2 - c_{s1}),$$

well defined because $\beta^2 - c_{s1} \geq 0$. The function $\hat{c}_{s2}$ is continuous and nondecreasing by properties (III) and (IV) above of q.

If $\beta^3 - \hat{c}_{s2}\ (c_{s1}) < 0$, then the initial condition c_{s1} generates the finite sequence of mature-age consumptions $(\hat{c}_{s1}\ (c_{s1}),\ \hat{c}_{s2}\ (c_{s1}))$, and we write $T(c_{s1}) = 2$, i.e., the iteration stops at $\hat{c}_{s2}\ (c_{s1})$.

If $\beta^3 - c_{s2} \geq 0$, then define

$$\hat{c}_{s3}(c_{s1}) \equiv q(\beta^3 - \hat{c}_{s2}(c_{s1})),$$

where, again, $\hat{c}_{s3}$ is continuous and nondecreasing. If $\beta^4 - \hat{c}_{s3}(c_{s1}) < 0$, then the initial condition c_{s1} generates the finite sequence of mature-age consumptions $(\hat{c}_{s1}(c_{s1}), \hat{c}_{s2}(c_{s1}), \hat{c}_{s3}(c_{s1}))$, and we write $T(c_{s1}) = 3$, i.e., the iteration stops at $\hat{c}_{s3}(c_{s1})$. If, on the contrary, $\beta^4 - \hat{c}_{s3}(c_{s1}) \geq 0$, then the iteration continues.

Iteratively, if $\beta^\tau - c_{s,\tau-1} \geq 0$, for $\tau = 2, \ldots, t$, then define

$$\hat{c}_{st}(c_{s1}) \equiv q(\beta^t - \hat{c}_{s,t-1}(c_{s1})),$$

where $\hat{c}_{st}$ is continuous and nondecreasing.

If, for some t, $\beta^\tau - c_{s,\tau-1} \geq 0$, $\tau = 2, \ldots, t$, but $\beta^{t+1} - \hat{c}_{st}(c_{s1}) < 0$, then the initial condition c_{s1} generates the finite sequence of mature-age consumptions $(\hat{c}_{s1}(c_{s1}), \hat{c}_{s2}(c_{s1}), \ldots, \hat{c}_{st}(c_{s1}))$, and we write $T(c_{s1}) = t$, i.e., the iteration stops at $\hat{c}_{st}(c_{s1})$.

If, on the contrary, $\beta^{\tau} - c_{s,\tau-1} \geq 0$, for all $\tau \geq 2$, then the initial condition c_{s1} generates the infinite sequence of mature-age consumptions $(\hat{c}_{s1}(c_{s1}), \hat{c}_{s2}(c_{s1}), \ldots, \hat{c}_{st}(c_{s1}), \ldots..)$, and we write $T(c_{s1}) = \infty$. In particular, $T(0) = \infty$, because if $c_{s1} = 0$, then $\Lambda(\beta - c_{s0}, c_{s1}) = \Lambda(\beta - c_{s0}, 0) < \Lambda(\beta^2, 0) = \Lambda(\beta^2 - c_{s1}, 0)$, and the law of motion gives $\hat{c}_{s2}(0) = 0$, which in turn implies that $\Lambda(\beta - c_{s0}, c_{s1}) < \Lambda(\beta^3, 0) = \Lambda(\beta^3 - c_{s2}, 0)$, and so on.

Note that $T(c_{s1}) \geq 2, \forall c_{s1} \in [0, \beta^2]$. Because the functions $\hat{c}_{st}$ are nondecreasing, $\overline{c}_{s1} < \overline{\overline{c}}_{s1}$ implies that $T(\overline{c}_{s1}) \geq T(\overline{\overline{c}}_{s1})$.

In words, given its maturity consumption $c_{mt} = \beta^t - c_{s,t-1}$, c_{st} is set (whenever possible) at the level that gives Generation t exactly the standard of living, Λ_1, of Generation One as determined by the initial conditions c_{s0} and c_{s1}.

Consider the set $\hat{C} = \{c_{s1} \in [0, \beta^2] \; : \; \exists t \in \{1, \ldots, T(c_{s1})\}$ such that $\beta^{t+1} - \hat{c}_{st}(c_{s1}) = 0\}$. This set is nonempty (because $\beta^2 - \hat{c}_{s1}(\beta^2) = 0$, i.e., $\beta^2 \in \hat{C}$). For $c_{s1} \in \hat{C}$, define $\hat{T}(c_{s1})$ as the lowest positive integer t for which $\beta^{t+1} - \hat{c}_{st}(c_{s1}) = 0$, i. e.,

$$\beta^{\hat{T}(c_{s1})+1} - \hat{c}_{s,\hat{T}(c_{s1})}(c_{s1}) = 0, \tag{14.3.1}$$

and

$$\beta^{\tau+1} - \hat{c}_{s\tau}(c_{s1}) > 0, \forall \tau \in \{1, \hat{T}(c_{s1}) - 1\}. \tag{14.3.2}$$

Note that, if we choose $\hat{t}$ such that $\beta^{\hat{t}} > c'$, we have that $\hat{T}(c_{s1}) \leq \hat{t}$, because $\hat{c}_{st}(c_{s1}) \leq c', \forall c_{s1}, \forall t$. Hence, the set $T^* \equiv \{t \geq 1 : \exists c_{s1} \in \hat{C}$ such that $t = \hat{T}(c_{s1})\}$ is a subset of the finite set $\{1, \ldots, \hat{t}\}$, and therefore there exists a $t^* \in T^*$ such $t > t^*, \forall t \in T^*, t \neq t^*$. Choose c^* in the nonempty set $\{c_{s1} \in [0, \beta^2] : \hat{T}(c_{s1}) = t^*\}$.

We then have that $\beta^{\tau+1} - \hat{c}_{s\tau}(c^*) > 0, \forall \tau < t^*$ (by (14.3.2)), and hence that $\Lambda(\beta^{\tau} - \hat{c}_{s,\tau-1}(c^*), \hat{c}_{s\tau}(c^*)) = \Lambda_1, \forall \tau < t^*$, i .e., all generations before t^* reach the standard of living Λ_1. Moreover, because, by (14.3.1), $\beta^{t^*+1} - \hat{c}_{st^*}(c^*) = 0$, we have that $\hat{c}_{st^*}(c^*) > 0$ and, thus, the standard of living of Generation t^*, $\Lambda(\beta^{t^*} - \hat{c}_{s,t^*-1}(c^*), \hat{c}_{st^*}(c^*))$, is also Λ_1.

Because $\beta^{t^*+1} - \hat{c}_{st^*}(c^*) = 0 \geq 0$, $t^* + 1 \leq T(c^*)$ and hence $\hat{c}_{s,t^*+1}(c^*)$ is well defined. If $\hat{c}_{s,t^*+1}(c^*) > \beta^{t^*+2}$, then by appealing to the continuity of the solution functions $\hat{c}_{st}(c_{s1})$ and the fact that $\hat{c}_{st}(0) = 0$, one can prove the existence of a $\overline{c}$ such that $\hat{c}_{s,t^*+1}(\overline{c}) = \beta^{t^*+2}$ and $\hat{c}_{st}(\overline{c}) < \beta^{t+1}$, for all $t < t^* + 1$, yielding $t^* + 1 \in T^*$, which contradicts the fact that t^* is the maximal element of T^*. It follows that $\hat{c}_{s,t^*+1}(c^*) \leq \beta^{t^*+2}$, which in turn implies that

$$\Lambda_1 \leq \Lambda(\beta^{t^*+1} - \hat{c}_{st^*}(c^*), \hat{c}_{s,t^*+1}(c*)) = \Lambda(0, \hat{c}_{s,t^*+1}(c^*)) \leq \Lambda(0, \beta^{t^*+2}).$$

It can then be checked that the leximin path is

$$\{c_{s0}, (\beta - c_{s0}, \hat{c}_{s1}(c^*)), (\beta^2 - \hat{c}_{s1}(c^*), \hat{c}_{s2}(c^*)), \ldots,$$
$$(\beta^{t^*} - \hat{c}_{s,t^*-1}(c^*), \hat{c}_{st^*}(c^*)), (0, \beta^{t^*+2}), (0, \beta^{t^*+3}), \ldots.\},$$

i .e., for $t=1,\ldots,t^*-1$, Generation t's standard of living is $\Lambda(\beta^t-\hat{c}_{s,t-1}(c^*), \hat{c}_{st}(c^*))=\Lambda_1$, that of Generation t^* is $\Lambda(\beta^{t^*}-\hat{c}_{s,t^*-1}(c^*), \hat{c}_{st^*}(c^*))=\Lambda(\beta^{t^*}-\hat{c}_{s,t^*-1}(c^*), \beta^{t^*+1}))=\Lambda_1$, that of Generation t^*+1 is $\Lambda(0,\beta^{t^*+2})\geq\Lambda_1$ and for $\tau=2,3,\ldots,$, that of Generation $t^*+\tau$ is $\Lambda(0,\beta^{t^*+\tau+1})>\Lambda(0,\beta^{t^*+\tau})$. One can check that any feasible path that gives any generation in the set $\{1,\ldots,t^*\}$ a standard of living higher than Λ_1 must give some other generation in the same set a standard of living lower than Λ_1.

Intuitively, recall that, when $c_{s0}=\beta$, then the leximin path is the everything-to-seniors path $\{c_{s0},(0,\beta^2),(0,\beta^3),\ldots\}$, which incidentally satisfies the sufficientarianist constraints when $\sigma=1$, and, consequently, $\mu=0$. But, if $c_{s0}<\beta$, then the earlier generations can achieve higher standards of living than $\Lambda(0,\beta^{t+1})$ by getting some consumption at maturity and not exhausting output at seniority. This gives a flat segment in the standard-of-living path, along which an increasing fraction of output is transferred to seniors. At the switch period t^*+1, all output is transferred to the Generation t^*'s seniors, and the standards of living follow from then on the everything-to-seniors path, which is increasing in t, so that there is human development after t^*. In particular, there is weak human development understood as the nondecreasingness and nonconstancy of the standard-of-living path.

14.3.4 An illustration

We can use the example of section 14.3.3 above to compare, for illustration purposes, the various paths discussed. As in there, let $\beta=1.1$ and $\Lambda(c_m,c_s)=\sqrt{c_m}+\sqrt{c_s}$, and recall that the best sufficientarianism involves $\sigma^*\equiv\frac{\beta}{1+\beta}=0.5238$.

Figure 14.3.2 depicts the fraction of (period-t)-output being transferred back to period-t seniors (i.e., the seniors of Generation $t-1$) along three paths: the best-sufficientarianist path, the 'everything to the seniors' path, and the leximin path with the same initial condition as the best sufficientarianist (i.e., with $c_{s0}=\sigma^*\beta$). We see that, in the latter, the fraction starts at σ^*, as in the first one, but it then raises to 100 per cent, which is reached at period $t^*+1=10$, and, after that, all output keeps going to the seniors.

Figure 14.3.3 compares the three standard-of-living paths. We observe:

- The best sufficientarianist standard-of-living path is always above the 'everything to the seniors' path: this is not surprising, since the latter has an implicit obligation to the past (namely, $c_{s0}=\beta$) much higher than that of the best sufficientarianist (namely, $\sigma^*\beta$).
- As expected, the leximin path with the same initial obligation as the best sufficientarianist path improves the utility of the worst-off generation (Generation One) relative to the best sufficientarianist path. But in our numerical example only two generations do better in the leximin path than in the best sufficientarianist. From Generation Three on, everybody is doing better in the best sufficientarianist path than in the leximin path.

This is only a numerical example, but it suggests that 'good' sufficientarianist paths may entail a relative small sacrifice earlier on, at substantial later benefits, than the leximin path with the same initial obligation to date-one seniors.

14.4 Consumption–investment trade-offs and human development in leximin

14.4.1 Leximin in OLG economies with capital accumulation

The only issue in section 14.3 was how to allocate among generations a good the output of which increased over time. As we saw, depending on the initial obligations towards date-one seniors, the leximin solution, perhaps after an early interval of constant standards of living, implied a strictly increasing path with all output going to the seniors. Indeed, if Generation $t+1$ is better off than Generation t along a nondecreasing leximin path of living standards, the amount of date-t good allocated to Generation t must be maximized, perhaps subject to the constraints discussed in section 14.2 above.

The argument was carried out under the simplifying assumption of a single, nonproduced input of production, namely labour. But if produced goods are also inputs in production, then a new question must be asked, namely how to allocate output between consumption and investment, in addition to the issue of how to allocate consumption among generations. As before, if the leximin path implies nondecreasing standards of living, and if Generation $t+1$ is better off than Generation t along such a path, then the amount of date-t good allocated to Generation t must be maximized given the values of all past variables. This in particular implies choosing date-t variables in order to maximize the amount of consumption good produced, to be made available to Generation t. Depending on the data in the economy, the maximization of date-t consumption given past variables may imply lower consumption levels in the future, in which case leximin will be incompatible with strictly increasing standards of living.

This section considers simple OLG economies with produced capital and poses the following question. Suppose that, for $t = 1, 2, \ldots$ aggregate consumption is maximized given past variables. Does this result in an increasing consumption path? If no, then leximin is incompatible with human development. If yes, then much of the intuition developed in the previous section on leximin paths involving human development can in principle be applied to this more realistic setting.

14.4.2 The conventional wisdom of the Diamond model

We adopt the well-known model of overlapping generations with capital due to Peter Diamond (1965). The temporal structure and the consumption sector are the same as in the OLG economy of section 14.3.1 above. But, instead of being produced by a single input with exogenously increasing productivity, output is produced according to the production function $F(K_t, L_t)$, where K_t and L_t are, respectively, the mean operating capital and mean labour employed in period t. A datum in the model is the amount of capital $\hat{K}_1$ that period 1 inherits from the past.

Later in this section we distinguish between the amount of capital *employed* in period t, K_t, and the amount of capital that period t *inherits* from the past, $\hat{K}_t$, although in Diamond's (1965) model they are identical. They obey the law of motion of capital

$$K_1 = \hat{K}_1, \hat{K}_1 \text{ is a datum,} \tag{14.4.1a}$$

$$\hat{K}_{t+1} = K_{t+1} = (1-\delta)K_t + I_t, \quad t = 1, 2, \ldots, \tag{14.4.1b}$$

$$I_t = F(K_t, L_t) - C_t, \quad t = 1, 2, \ldots, \tag{14.4.1c}$$

where C_t denotes the amount of output made available for consumption in period t, I_t is period t's new investment, and $\delta \in [0, 1]$ is a given depreciation rate.

The system of equations (4.1a–c) can be rewritten

$$K_1 = \hat{K}_1, \hat{K}_1 \text{ is a datum,} \tag{14.4.1a}$$

$$K_t = (1-\delta)K_{t-1} + I_{t-1}, \quad t = 2, 3, \ldots, \tag{14.4.1d}$$

$$C_t = F(K_t, L_t) - I_t, \quad t = 1, 2, \ldots. \tag{14.4.1c}$$

Again, we postulate certainty, a representative consumer in each generation, and a fixed and constant labour supply in period t, normalized to $L_t = 1$, $\forall t \geq 1$. Figure 14.4.1 reproduces Figure 14.3.1 above, but with the new description of the amount of the good produced in each period.

The law of motion (14.4.1) is interpreted as follows: new capital goods become operational in the period that follows the period of their production, an initial condition being the amount $\hat{K}_1$ that period one inherits from the past, which determines the amount of capital employed in period one as $K_1 = \hat{K}_1$. Capital depreciates while employed, so that the mean amount of capital K_{t+1} operational in period $t+1$ is the sum of the depreciated old capital, $(1-\delta)\hat{K}_t$, and period t's new investment I_t.[7]

As in section 14.3.2 above, the standard of living of Generation $t (t \geq 1)$ is a function $\Lambda(c_{mt}, c_{st})$ of its consumption when mature, c_{mt}, and its consumption when senior, c_{st}, defined on $\Re^2_+$.[8] The physical balance condition for consumption is $C_t = c_{mt} + c_{s,t-1}$.

The first decision period is $t = 1$, where a senior Generation Zero is alive. As in section 14.3.4 above, because Generation Zero is never mature, its standard

	Period 1 Capital Inherited $= \hat{K}_1$ Output $= F(K_1, L_1)$	*Period 2* Capital Inherited $= \hat{K}_2$ Output $= F(K_2, L_2)$	*Period 3* Capital Inherited $= \hat{K}_3$ Output $= F(K_3, L_3)$		. . .	*Period t* Capital Inherited $= \hat{K}_t$ Output $= \beta^t$	*Period* $t+1$ Capital Inherited $= \hat{K}_{t+1}$ Output $= F(K_{t+1}, L_{t+1})$	. . .
Gen. 0	Senior c_{s0}							
Gen. 1	Mature c_{m1}	Senior c_{s1}						
Gen. 2		Mature c_{m2}	Senior c_{s2}					
Gen. 3			Mature c_{m3}	Senior c_{s3}				
. . .				. . .	. . .	. . .		
Gen. t						Mature c_{mt}	Senior c_{st}	
. . .							. . .	. . .

Figure 14.4.1 Diamond's (1965) model
Each cell displays the notation for the corresponding amount of consumption

of living is not defined and we postulate that Generation Zero is entitled to a given amount c_{s0} of the consumption good produced in period one. Accordingly, the initial conditions are $(c_{s0}, \hat{K}_1)$, assumed to satisfy $c_{s0} \leq F(\hat{K}_1, 1)$.

Diamond (1965) allows for negative investment, on the grounds that 'since capital and output are the same commodity, one can consume one's capital' (p. 1127). Equality (14.4.1) implies in this case that one unit of capital can be converted back into $(1 - \delta)$ units of consumption.

This technology does not allow for human development, even understood in the weak sense of a nondecreasing and nonconstant path of standards of living, at a maximin solution. Indeed, assume on the contrary that a maximin path $(\Lambda_1, \Lambda_2, \ldots)$ of standards of living satisfies

$$\Lambda_1 = \Lambda_2 = \cdots = \Lambda_t < \Lambda_{t+1} \leq \Lambda_{t+2} \leq \cdots$$

for some $t \geq 1$. Then Generation t is a worst-off generation. But, as long as Λ is increasing in c_{st}, this would imply that $I_t = 0$, as in point A of Figure 14.4.2: otherwise, because capital does not enter the standard-of-living function, one could increase Generation t's standard of living by a small reduction in I_t still allowing for Generations $t + \tau$, $\tau = 1, 2, \ldots$, a standard of living not lower than Λ_t . Similarly, we must have $c_{st} = C_t$ and hence $c_{m,t+1} = 0$, so that $\Lambda_{t+1} = \Lambda(0, F(0,1))$. But the path $(c_{s0}, (0, F(0,1))$ is certainly feasible: thus, the value of the maximin problem, namely Λ_t, must be no lower than it, i.e., $\Lambda(0, F(0,1)) \leq \Lambda_t$, contradicting the fact that $\Lambda_t < \Lambda_{t+1} = \Lambda(0, F(0,1))$. Thus, no human development, even in the weak sense, is possible at a maximin solution of this economy.

Note that the fact that capital becomes productive only one period after being produced plays a central role in the argument.

The assumption that capital goods can be costlessly recycled into consumption is unrealistic for produced capital,[9] but in any event it is not crucial for the argument. Indeed, one may add to (14.4.1) the condition that investment I_t cannot be negative, i.e., capital is of the putty-clay variety, and cannot be converted into a consumption good after it has been produced. Formally,

$$I_t \equiv F(K_t, L_t) - C_t \geq 0, \tag{14.4.2}$$

which, recalling (14.4.1), implies,

$$\hat{K}_{t+1} \geq (1 - \delta)\hat{K}_t, \tag{14.4.3}$$

an inequality which will be maintained in what follows. But human development would be problematic in a maximin solution even then. Whether (14.4.2) is imposed or not, (14.4.1) implies a (one-for-one) trade-off between current consumption and the amount of capital bequeathed to the following period (which determines the productivity of its labour), because, by (14.4.1), $C_t = F(\hat{K}_t, 1) - I_t$, so that for given L_t and K_t, $\frac{dC_t}{dI_t} = -1$ as depicted in Figure 14.4.2. Under (14.4.2), given $K_t = \hat{K}_t$ (inherited from the past) and $L_t = 1$, the maximization of the amount of the consumption good available in period t requires $F(K_t, L_t) = C_t$, as in point B of Figure 14.4.2, and hence it requires to push I_t and $\hat{K}_{t+1}$ down to their minimum values 0 and $(1-\delta)\hat{K}_t \leq \hat{K}_t$, respectively.

Of course, social or institutional constraints could impose a positive lower bound on I_t, in which case (14.4.2) would be replaced by

$$F(K_t, L_t) - C_t \geq m_t. \qquad (14.4.4)$$

Maximizing C_t under (14.4.1) and (14.4.4) would then yield

$$\hat{K}_{t+1} = F(\hat{K}_t, L_t) - C_t + (1-\delta)\hat{K}_t = m_t + (1-\delta)\hat{K}_t,$$

which is greater than $\hat{K}_t$ if m_t is large relative to $\delta \hat{K}_t$. The appendix below applies this idea to the Roemer–Veneziani (2004) model, but in this section we maintain (14.4.2).

What would be the paths of the variables if, at any period, aggregate consumption were maximized given the variables determined in previous dates? In period one, output is $F(\hat{K}_1, 1)$, determined by the initial condition $\hat{K}_1$, and the maximization of C_1 given $\hat{K}_1$ yields $I_1 = 0$. Therefore, $\hat{K}_2 = (1-\delta)\hat{K}_1$, requiring $I_1 = 0$ if C_2 is maximized given $\hat{K}_2$, and so on. Thus, the time path of investments is $(I_1, I_2, \ldots) = (0, 0, \ldots.)$, and that of consumption is $(C_1, C_2, \ldots) = (F(\hat{K}_1, 1)$, F($(1-\delta)\hat{K}_1$,1), ...), decreasing (resp. constant) if $\delta > 0$ (resp. $\delta = 0$). Thus, as argued in section 14.4.1 above, leximin is incompatible with human development in this economy.

14.4.3 Investment productive within the period

Of course, both 'consumption' and 'capital' are aggregates. Some activities have both consumption and investment components: the development literature emphasizes this dual role in nutrition, education and health services (Gersovitz, 1988; Steger, 2002): a well-nourished, educated and healthy workforce is more productive than a malnourished, illiterate and ailing one, voiding or mitigating the just discussed tradeoff between current consumption and future labour productivity. But this idea does not apply to our trade-off between the consumption of the old generation at the end of its life, when it is no longer working, and the productivity of the labour of the younger generation.

Yet the stark conflict between consumption and investment depicted in Figure 14.4.2 has unrealistic features, even in the case where all consumption

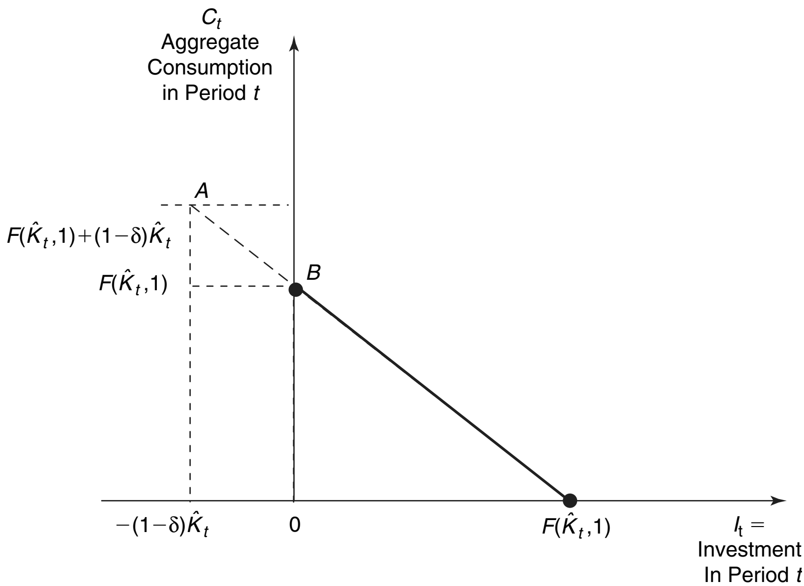

Figure 14.4.2 The consumption–investment conflict in Diamond's model

goes to senior people with no productive future. As noted in section 14.3.2 above, here a period is interpreted as entailing a relatively long time span, say more than twenty years. It is unreasonable to assume that the amount of capital in operation throughout that time is constant at the level inherited at the beginning of the period. In fact, the production of most consumption goods requires forms of investment that increase the capital stock not only in the future, as in the Diamond model, but also in the current period. The amount of the good available for consumption during this interval may conceivably be increased by diverting some period-t resources to the production of capital goods that are useful within period t, in addition to increasing the amount of capital available at the beginning of period $t+1$. Figure 14.4.2, on the contrary, shows consumption within period t being maximized in Diamond's framework when I_t is pushed down to zero.

A rigorous development of this idea may require subdividing period t, the period during which Generation t is mature and Generation $t+1$ senior, into a number of shorter time intervals, distinguishing the production activities among them, and considering disaggregated capital goods with different production lags and depreciation rates. For the sake of simplicity, however, let us assume that the amount $K_t$of capital operational in period t is a function $\chi(\hat{K}_t, I_t)$ of the amount of capital that period t inherits from the past, $\hat{K}_t$, and period t's output I_t of capital goods, with $\frac{\partial \chi}{\partial I_t} \geq 0$. As before, period t's total output is the sum of period t's consumption-goods output, C_t, and period t's capital-goods output I_t, i.e., $I_t = F(K_t, L_t) - C_t$. The law of motion (14.4.1) now generalizes to

$$K_t = \chi(\hat{K}_t, I_t), t = 1, 2, \ldots, \hat{K}_1 \text{ is a datum,} \tag{14.4.5a}$$

$$\hat{K}_{t+1} = (1 - \delta)K_t + \eta I_t, t = 1, 2, \ldots, \tag{14.4.5b}$$

$$I_t = F(K_t, L_t) - C_t, t = 1, 2, \ldots, \tag{14.4.5c}$$

which is exactly (14.4.1) in the special case where $\chi(\hat{K}_t, I_t) = \hat{K}_t$ and $\eta = 1$. Now we allow for $\eta \in [0, 1]$ with the interpretation that we may wish to avoid double-counting those capital goods produced during period t that became operational in period t (already embodied in K_t), as well as to capture possible depreciation.

We maintain the assumption that $I_t \geq 0$. Moreover, if $\chi(\hat{K}_t, I_t) \geq \hat{K}_t$ (reflecting the putty-clay character of capital), then

$$\hat{K}_{t+1} \geq (1 - \delta)K_t = (1 - \delta)\chi(\hat{K}_t, I_t) \geq (1 - \delta)\hat{K}_t, \quad t = 1, 2, \ldots$$

as in (14.4.3) above.

Substituting (14.4.5a) into (14.4.5c), and letting $L_t = 1$, we obtain

$$C_t = F(\chi(\hat{K}_t, I_t), 1) - I_t, \quad t = 1, 2, \ldots,$$

The FOC for the maximization of C_t, given $\hat{K}_t$ and subject to $I_t \geq 0$, is now

$$\frac{dC_t}{dI_t} = \frac{\partial F}{\partial K_t}(\chi(\hat{K}_t, I_t), 1)\frac{\partial \chi}{\partial I_t}(\hat{K}_t, I_t) - 1 \leq 0, \tag{14.4.6}$$

with equality if $I_t > 0$. In Diamond's case, $\frac{\partial \chi}{\partial I_t} = 0$. Thus, $\frac{dC_t}{dI_t} = -1 < 0$, and C_t is maximized at $I_t = 0$, as noted. But if $\frac{\partial \chi}{\partial I_t} > 0$, then $\frac{dC_t}{dI_t} > 0$ whenever $\frac{\partial \chi}{\partial I_t}\frac{\partial F}{\partial K_t} > 1$, softening the consumption-capital tradeoff, which may occur if F is strictly concave and K_t low.

Figure 14.4.3 provides a simple illustration for the following example. Let $\delta = 0.1$, $\eta = 0$, $F(K_t, L_t) = 10(K_t)^{3/4}(L_t)^{1/4}$, $\chi(\hat{K}_t, I_t) = \hat{K}_t + 0.9I_t$ and let $\hat{K}_t = 200$. Note the contrast with Figure 14.4.2: in order to maximize consumption in period t given $\hat{K}_t$, investment in period t should be set at $2084.38 > 0$. As a result, the amount of capital that period t bequeaths to period $t+1$ is now $\hat{K}_{t+1} = 1868.35$, substantially larger than the amount of capital $\hat{K}_t = 200$ that period t inherited from the past and, *a fortiori*, of the lower bound $(1 - \delta)\hat{K}_t = 180$ for $\hat{K}_{t+1}$.

Summarizing, when investment is at least partly operational in the period where it is produced, it may very well entail positive investment when

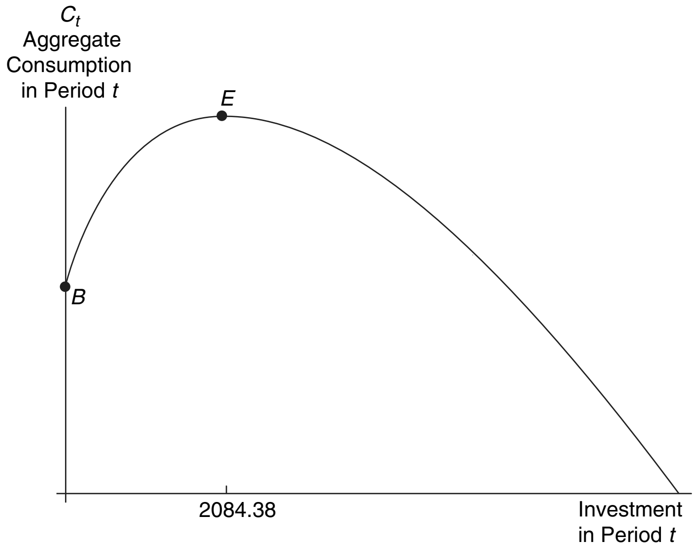

Figure 14.4.3 The consumption–investment frontier when investment is productive within the period

investment is chosen in order to maximize the period's consumption, given past variables. Can the sequence of such maximization problems yield an ever-increasing path of consumption in the model of this subsection? We address this issue for the special case where the function $\chi(\hat{K}_t, I_t)$ is linear, i.e., $\chi(\hat{K}_t, I_t) = \rho\hat{K}_t + \psi I_t$, which, from (14.4.5a–b), implies:

$$K_t = \rho\hat{K}_t + \psi\frac{\hat{K}_{t+1} - (1-\delta)K_t}{\eta},$$

or

$$K_t = \frac{\eta\rho\hat{K}_t + \psi\hat{K}_{t+1}}{\eta + \psi(1-\delta)}.$$

We restrict ourselves to the case where (14.4.6) is always satisfied with equality, becoming

$$\frac{\partial F}{\partial K_t}\left(\frac{\eta\rho\hat{K}_t + \psi\hat{K}_{t+1}}{\eta + \psi(1-\delta)}, 1\right)\psi - 1 = 0. \qquad (14.4.7)$$

Equality (14.4.7) implicitly defines $\hat{K}_{t+1}$ as a function of $\hat{K}_t$, the graph of which is the phase line in the phase diagram with $\hat{K}_t$ on the horizontal axis and $\hat{K}_{t+1}$ on the vertical one. The implicit differentiation of (14.4.7) yields:

$$\frac{d\hat{K}_{t+1}}{d\hat{K}_t} = -\frac{\frac{\partial^2 F}{\partial K_t^2}\frac{\psi}{\eta+\psi(1-\delta)}\eta\rho}{\frac{\partial^2 F}{\partial K_t^2}\frac{\psi}{\eta+\psi(1-\delta)}\frac{\partial^2 F}{\partial K_t^2}\frac{\psi}{\eta+\psi(1-\delta)}\psi} = -\frac{\eta\rho}{\psi} < 0.$$

Thus, the phase line crosses the 45° line only once, yielding a unique steady state $\overline{K}$, and the paths of $\hat{K}_t$ are oscillatory, with oscillations of decreasing (resp. increasing, resp. constant) amplitude if $\frac{\eta\rho}{\psi} < 1$ (resp., $\frac{\eta\rho}{\psi} > 1$, resp. $\frac{\eta\rho}{\psi} = 1$). In any case, the path of $\hat{K}_t$ fails to be monotonically increasing, implying that the path of C_t is not monotonically increasing, because C_t is increasing in $\hat{K}_t$. (A larger $\hat{K}_t$ relaxes the constraint in the problem of choosing I_t in order to maximize C_t subject to $I_t + C_t - F(\chi(\hat{K}_t, I_t), 1) \leq 0$.) Thus, if C_t is maximized in each period given past variables, then the resulting path is not increasing. It follows that leximin is incompatible with human development in this case.

14.4.4 Endogenous growth

It has long been recognized that production and investment generate technological progress that may spill over the boundaries of individual firms, and this idea has led to the now-dominant paradigm of endogenous growth.[10] In agreement with this tradition, I now postulate a third input of production, namely the stock of knowledge A_t, a stock that grows with investment, output or level of capital. Again, I notationally distinguish between A_t, the stock of knowledge operational in period t, and $\hat{A}_t$, the stock of knowledge that period t has inherited from the past. For simplicity, I assume that the improvement in knowledge or learning occurs as a side effect of producing new capital, as proposed by Arrow (1962). The combination of the productivity of new capital within the period (section 14.4.3) and the effect of new investment in the future stock of knowledge yields the following laws of motion:

$$K_t = \chi(\hat{K}_t, I_t), \;\; t = 1, 2, \ldots, \hat{K}_1 \text{ is a datum,} \tag{14.4.8a}$$

$$\hat{K}_{t+1} = (1-\delta)K_t + \eta I_t, \;\; t = 1, 2, \ldots, \tag{14.4.8b}$$

$$A_t = \omega(\hat{A}_t, I_t), \;\; t = 1, 2, \ldots, \hat{A}_1 \text{ is a datum,} \tag{14.4.8c}$$

$$\hat{A}_{t+1} = (1-\delta_A)A_t + \nu(I_t), \;\; t = 1, 2, \ldots, \tag{14.4.8d}$$

$$I_t = f(A_t, K_t, L_t) - C_t, \;\; t = 1, 2, \ldots, \tag{14.4.8e}$$

where f is the production function, the functions ω and ν describe the contribution of new investment to the accumulation of knowledge, δ_A is the obsolescence rate of knowledge, and, for $t \geq 1$, A_t is the stock of knowledge that is operational in period t. Otherwise, the model is as in section 14.4.2, and in particular we impose the condition $I_t \geq 0$.

It turns out that now, depending on the various functional forms and parameters, the sequential maximization of aggregate consumption given past variables may generate an increasing path of consumption, and hence the leximin solution may be compatible with human development.

In order to obtain an easily computable solution, we specialize (14.4.8) as follows. First, set $\chi(\hat{K}_t, I_t) = \hat{K}_t + \psi I_t$, $\omega(\hat{A}_t, I_t) = \hat{A}_t$, (i.e., $A_t = \hat{A}_t$: knowledge is not productively disseminated until after the period in which it is created), $\delta_A = 1$ (knowledge becomes obsolete after one period), $f(A_t, K_t, L_t) = A_t K_t^{\alpha} L_t^{1-\alpha}$, $\alpha \in (0,1)$, and $\nu(I_{t-1}) = \xi(I_{t-1})^{\gamma}$, $\xi > 0, \gamma > 0$. Second, replace (14.4.8a,b) by (14.4.9a,b) below (i.e., $\delta = \eta = 1$, and χ is a linear function: the interpretation is that new capital becomes obsolete after two periods). The model then specializes to

$$K_t = \lambda \hat{K}_t + \psi I_t,\ t = 1, 2, \ldots, \hat{K}_1 \text{ a datum,} \tag{14.4.9a}$$

$$\hat{K}_{t+1} = I_t,\ t = 1, 2, \ldots, \tag{14.4.9b}$$

$$A_t = \xi(I_{t-1})^{\gamma},\ t = 1, 2, \ldots,$$

$$C_t = A_t K_t^{\alpha} L_t^{1-\alpha} - I_t,\ t = 1, 2, \ldots,$$

with all parameters positive. For $t = 1, 2\ldots$, the programme of the maximization of C_t given given I_{t-1} is

$$\text{Max}_{I_t \geq 0} \xi(I_{t-1})^{\gamma}[\lambda I_{t-1} + \psi I_t]^{\alpha} - I_t, \tag{14.4.10}$$

where $I_0 \equiv \hat{K}_1$. The first-order condition (parallel to (14.4.6)) is now

$$\xi(I_{t-1})^{\gamma} \alpha [\lambda I_{t-1} + \psi I_t]^{\alpha - 1} \psi \leq 1, \tag{14.4.11}$$

with equality if $I_t > 0$, which occurs if and only if

$$\xi(I_{t-1})^{\gamma} \alpha [\lambda I_{t-1}]^{\alpha - 1} \psi > 1, \tag{14.4.12}$$

in which case the equality in (14.4.11) yields

$$I_t = \Phi(I_{t-1}) \equiv (\xi\alpha)^{\frac{1}{1-\alpha}} \psi^{\frac{\alpha}{1-\alpha}} I_{t-1}^{\frac{\gamma}{1-\alpha}} - \frac{\lambda}{\psi} I_{t-1} \equiv a I_{t-1}^{\frac{\gamma}{1-\alpha}} - b I_{t-1}, \tag{14.4.13}$$

where $a \equiv (\xi\alpha)^{\frac{1}{1-\alpha}} \psi^{\frac{\alpha}{1-\alpha}} > 0$ and $b \equiv \frac{\lambda}{\psi} > 0$. Equation (14.4.13) gives the phase line for the phase diagram with I_{t-1} on the horizontal axis and I_t on the vertical axis.

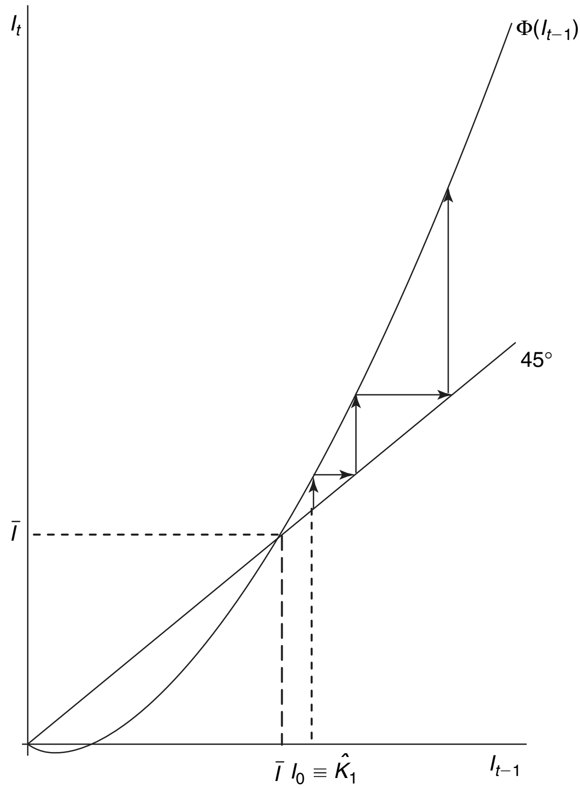

Figure 14.4.4 Progress at the leximin of the endogenous growth model with same-period productive investment

Assume that $\gamma > 1 - \alpha$. Then Φ is strictly convex and has a positive fixed point $\bar{I} \equiv \left(\frac{1+b}{a}\right)^{\frac{1-\alpha}{\alpha+\gamma-1}}$ (in addition to the zero fixed point), with $\Phi(I_{t-1}) > I_{t-1}$ whenever $I_{t-1} > \bar{I}$. Moreover, it is easy to check that, if $I_{t-1} > \bar{I}$, then (14.4.12) holds, so that (14.4.11) is indeed satisfied with equality, and (14.4.13) applies. Hence, if $\hat{K}_1 > \bar{I}$, then I_t follows an increasing path, as in Figure 14.4.4. Therefore, so does C_t, as it immediately follows from applying the envelope theorem to the objective function in (14.4.10).

Summarizing, if the economy is sufficiently productive ($\alpha + \gamma > 1$), and the initial level $\hat{K}_1$ of capital is sufficiently high, then the sequential maximization of consumption given past variables yields an increasing path of consumption, reminiscent of the paths assumed in section 14.3 above, and leximin is in principle compatible with human development.

14.5 Leximin-sufficientarianism complementarities

I argue that the two approaches must be to some extent complementary in the intergenerational application. On the one hand, the sensible implementation of the leximin criterion calls for socially-determined bounds close in spirit to the sufficientarianist floors. On the other, the sufficientarianist criterion may need to be complemented by an objective function.

14.5.1 Social constraints in the leximin optimization

First, as discussed in section 14.3 above, the intergenerational implementation of leximin requires specifying the levels of primary goods for the generations that are old at the time of decision: these obligations to initial seniors are often justified in sufficiency terms.

Second, the irreversibility of certain forms of environmental depletion (as in the extinction of species), and the possibility of catastrophic effects of climate changes may justify the imposition of lower bounds on natural capital in the spirit of 'safe minimum standards' (Arrow and Fisher, 1974; Bishop, 1978).

Third, modern societies prohibit child labour and have compulsory education laws, hence imposing lower bounds on investment in education and the human-capital model with minimum education standards due to Roemer and Veneziani (2004) illustrates this idea: see the appendix below.

Fourth, the uncertainty about the time of death may impose upper bounds on backward transfers. The reasoning goes as follows. On the one hand, as noted in section 14.2.1 above, there is the wish to maximin (or leximin) the standards of living of different generations over their lifespans, rather than the contemporaneous 'standards of living' of the currently young, mature and old. If there is individual uncertainty within the individuals of a generation relative to the date of death, then it matters whether the lifetime standards of living are evaluated *ex ante*, when the generation is born, *ex post*, at the end of each person's life, or *ex interim*, as the uncertainty resolves. To the extent that the date of death may be independent of behaviour, the basic maximin logic of minimizing the dependence of a person's quality of life on circumstances beyond her control advocates an *ex interim* approach. But then we face basic noncomparabilities, first between the *ex post* welfare of individuals who die young and those who die old (other things being equal, what is better, dying young or living a long life but being destitute in the old years?) and also between the *ex interim* welfare, at date t, of earlier-generation individuals who have reached old age, and later generation individuals who are currently in their working years but may or may not reach old age. In any event, the uncertainty about the time of death may place natural upper bounds on the acceptable backward transfers from the latter to the former, which, in particular, may alleviate the 'everything-to-seniors' problem.

In summary, the implementation of leximin often appeals to socially-determined standards on some forms of transfers and investments that are ultimately justified by adequacy considerations, in the spirit of the sufficientarianist constraints.

Note that the first and fourth ideas above are based on uncertainty considerations, indicating the need to explicitly incorporate uncertainty in the analysis of intertemporal equity.

14.5.2 Objective functions and optimization in sufficientarianism

In the simple model of section 14.3.2, the evolution of the economy over time is sensitive to the specific sufficientarianist floors, and the introduction of an objective function may help select desirable paths. More specifically, each value of the sufficientarianist parameter σ, together with the minimization of backward transfers, determines there a path $\Xi(\sigma)$ for the economic variables relevant for the standard of living of the various generations. As noted there, these paths can be ranked by the Pareto criterion, so that there is an unambiguously optimal value σ^*, yielding a 'best sufficientarianist' path $\Xi(\sigma^*)$, which is actually the leximin optimum among the paths generated values for the parameter σ in [0, 1].

This is a special case of a phenomenon which may appear in a more general model allowing for endogenous growth, where the production possibilities may depend on the path followed by the distribution of output, and several sufficientarianist parameters.

Let $x \equiv (x_1, x_2, \ldots, x_t, \ldots)$ denote a path of economic variables, and let $Y \equiv (Y_1, Y_2, \ldots, Y_t, \ldots)$ denote a sequence of attainable production sets given the technology and productive resources. In the example of section 14.3, Y was exogenous, but more generally we can assume that Y depends on the path x through endogenous progress and capital accumulation, and let the mapping $Y = \Upsilon(x)$ capture this dependence. Suppose that σ is a vector of sufficientarianist parameters interpreted as 'good enough' levels relative to the technological and physical possibilities Y, and that, as in section 14.3, a unique $x = \varphi(\sigma, Y)$ path is determined by σ and Y. The equality '$x = \varphi(\sigma, \Upsilon(x))$' implicitly defines a relation $x = \Xi(\sigma)$, so that a 'best sufficientarianist' vector σ of parameters can be chosen according to some criterion. Hence, in the intergenerational context we may want to apply a criterion such as the leximin to select desirable values for the sufficientarianist floors.

A different type of problem may appear when the admissible set under a given sufficientarianism norm is empty. Suppose, for example, that the only physically feasible paths are A and B of Figure 14.5.1, and that the sufficientarianist floor for the standard of living is Λ^*, satisfied by neither A nor B. This is a relevant case in the intertemporal problem, because it may be impossible for the earlier generations to achieve 'good enough' standards of living, in which case sufficientarianism is silent. But paths A and B violate the sufficientarianist constraints in different respects, and we may want to appeal

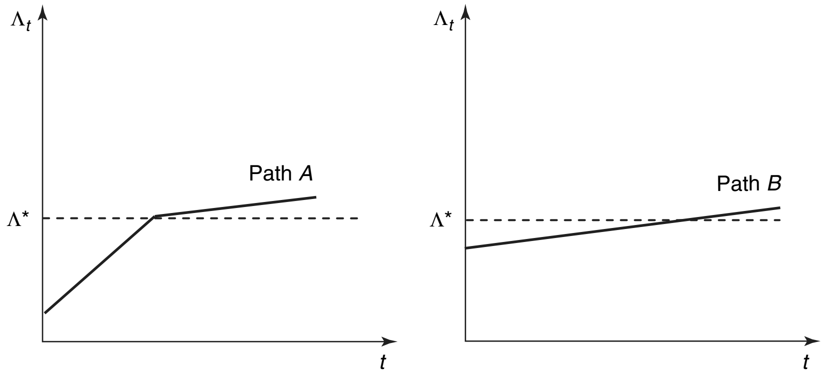

Figure 14.5.1 Two types of violation of the sufficientarianist constraint

to a criterion external to sufficientarianism. Of course, different criteria may choose different paths: in Figure 14.5.1, maximin chooses Path B, whereas utilitarianism favours Path A.

14.6 Concluding remarks

Two normative criteria have been considered: the Rawlsian maximin (or leximin), and the less demanding sufficientarianism. The discussion has taken place at a preliminary level, in extremely simple models, and should be considered as a first step for more realistic analysis. In particular, the study of sufficientarianism in models of endogenous growth has not been attempted.

A second open question is the incorporation of uncertainty in the standard-of-living function on which the leximin calculus is based. Third, environmental intergenerational externalities have not been explicitly considered.

Subject to these limitations, the discussion has argued the following points.

(1) *Adaptability of sufficientarianism to the intergenerational equity problem.* The sufficientarianism criterion presents some practical advantages over leximin in an intergenerational context, because it can be applied in a disaggregated manner, primary good by primary good, life stage by life stage, and state of the world by state of the world. Moreover, it is less biased against human development. But it faces the challenge, both in the atemporal and the intergenerational applications, of nonarbitrarily specifying 'good enough' norms

(2) *Central role of the initial obligations to seniors in the implementation of the leximin criterion.* Contrary to sufficientarianism, the initial obligations to decision-time seniors are exogenous in the leximin calculus.

On the one hand, they must be justified by appealling to 'sufficientarianist' considerations. On the other hand, the leximin optimal paths depend crucially on the level of these obligations. Thus, in order to make 'apples-to-apples' comparisons between the leximin and sufficientarianist paths, the initial obligations should be the same. One must note that non-overlapping generations models, or those models without 'senior' generations (i.e., models where forward transfers are perhaps possible, but backward transfers are not) are incapable of reflecting these issues.

(3) *Seriousness of the 'everything-to-seniors' problem in leximin.* Barring uncertainty, complementarities among consumption at different stages of life, satiety, and upper bounds on backward transfers (perhaps incentive-based or socially-determined), if there is human development at the leximin solution, then an extreme lifetime pattern results, with low levels of primary goods during the productive years, and abundance at the late stages of life. Again, models without 'senior' generations cannot convey the problem.

(4) *Consumption/investment conflicts in leximin may well be less serious.* On the other hand, because the production of consumption goods improves the technology, the traditionally emphasized tradeoff between current consumption and future productive capacities may not be severe enough to preclude human development at the leximin solution. This is ultimately an empirical question: a reliable answer will require the development of more complete, realistic models, with sufficient disaggregation to track the evolution of the various primary goods (health, education, the environment . . .) and to estimate empirically the relevant functional relations.

(5) *Leximin and sufficientarianism as complementary approaches.* When applied to the intergenerational problem, the sensible implementation of the leximin criterion calls for socially determined 'sufficientarianist' bounds on various forms of transfers and investment. Conversely, the path followed by the economy is sensitive to the specific sufficientarianist requirements, and the introduction of an objective function may help select desirable paths. Or the intertemporal sufficientarianist norms may be impossible to achieve for the earlier generations, in which case sufficientarianism is silent and decisions must appeal to an objective function.

Appendix: Lower bounds on investment in education

As noted in section 14.2, lower bounds on investment may yield progress at the maximin (or leximin) solution in an otherwise conventional model. This section develops this idea in the context of a recent model of investment in education due to Roemer and Veneziani (2004). Again two generations coexist, but, contrary to Figures 14.3.1 and 14.4.1 above, the productive age

(Mature) is older than the unproductive one (Young), see Figure A. Note that Generation t is young in period $t-1$.

People get educated when young, and work when mature. A person's living standard depends on her wage rate and on her consumption when mature.[11] Roemer and Veneziani 'imagine that a person's wage is a measure of her level of human capital, and individuals derive welfare directly from their human capital' via self-esteem and self-realization. A person's wage depends on the fraction of GNP spent on her education, financed by income taxes, and on the wage of her parents. Roemer and Veneziani consider two 'dynasties' differentiated by their initial wage levels. Here I simplify their model by assuming a single dynasty: i.e., there is a representative consumer in each generation, as in the preceding sections.

We take as initial condition the wage rate w_1, normalized to one, of Generation 1 (mature in period 1). The first decision period is again 1, and the standard of living of Generation 1 does enter the maximin calculus. The decision variables can be given as the sequence $(\pi_1, \pi_2, \ldots, \pi_t, \ldots)$ of income tax rates in $t = 1, 2, \ldots$. The law of motion of the wage rates is as follows: as just stated, w_1 is an initial condition. For period $t > 1$, $w_t = g(\pi_{t-1})h(w_{t-1})$ for some increasing functions g and h, where π_{t-1} and w_{t-1} are, respectively, the tax and wage rates in period $t-1$. For $t \geq 1$, given the tax rate π_t, the consumption of Generation t (in period t, when Generation t is mature) is $(1-\pi_t)w_t$. Let us postulate the Cobb-Douglas standard-of-living function $w_t^s[(1-\pi_t)w_t]^{1-s}$, and specialize the functions g and h to the linear $g(\pi_{t-1}) = \zeta\pi_{t-1}$ and $h(w_{t-1}) = \kappa w_{t-1}$, with $\zeta > 0$ and $\kappa > 0$.

First note that, as illustrated in Figure 14.4.2 for the Diamond model, there is a hard conflict between consumption and investment: a higher tax rate on generation t (paid in t, and devoted to educate the young Generation $t+1$) unambiguously decreases the standard of living of Generation t , and increases that of Generation $t+1$. Thus, as noted in section 14.2 above, unless a lower bound is imposed on π_t, maximin will be incompatible with progress.

But suppose now that compulsory education and child-labour laws impose a lower bound $\overline{\pi}$ on $\pi_t (t \geq 1)$. If this bound is high enough, and if the economy is productive enough, then the maximin solution implies progress: the relevant inequality turns out to be $\overline{\pi}\zeta\kappa > 1$. Indeed, in this case the maximization of the welfare of Generation 1 requires setting $\pi_1 = \overline{\pi}$, with standard of living $\Lambda^1 \equiv w_1^s(1-\overline{\pi})^{1-s} = (1-\overline{\pi})^{1-s}$, by the normalization. This yields $w_2 = (\zeta\overline{\pi})\kappa$, with Generation 2's standard of living equal to $\Lambda^2 \equiv (\zeta\overline{\pi}\kappa)^s(1-\pi_2)^{1-s} \geq (\zeta\overline{\pi}\kappa)^s(1-\overline{\pi})^{1-s}$, which is greater than Λ^1 whenever $\zeta\overline{\pi}\kappa > 1$. The maximization of the welfare of Generation 2 under the minimum education constraint yields again $\pi_2 = \overline{\pi}$, with $w_3 = (\zeta\overline{\pi})\kappa w_2 = (\zeta\overline{\pi})\kappa(\zeta\overline{\pi}\kappa) = (\zeta\overline{\pi}\kappa)^2$, and with Generation 3's standard of living equal to $\Lambda^3 \equiv (\zeta\overline{\pi}\kappa)^{2s}(1-\pi_3)^{1-s} \geq (\zeta\overline{\pi}\kappa)^{2s}(1-\overline{\pi})^{1-s}$. By iteration, we get

	Period 1	*Period 2*	*Period 3*	*Period 4*	...	*Period t*	*Period t*−1	...
Gen. 1	Mature w_1 $(1-\pi_1)w_1$							
Gen. 2	Young	Mature w_2 $(1-\pi_2)w_2$						
Gen. 3		Young	Mature w_3 $(1-\pi_3)w_3$					
Gen. 4			Young	Mature w_4 $(1-\pi_4)w_4$				
...				...	...	...		
Gen. $t+1$						Young	Mature w_t $(1-\pi_t)w_t$	
...							...	...

Figure 14.A Overlapping generations in the Roemer–Veneziani model
Each cell displays the notation for the wage rate and the amount of consumption

$\Lambda^t \equiv (\zeta\overline{\pi}\kappa)^{(t-1)s}(1 - \overline{\pi})^{1-s}$, i.e., $\frac{\Lambda^{t+1}}{\Lambda^t} = \frac{(\zeta\overline{\pi}\kappa)^{ts}(1-\overline{\pi})^{1-s}}{(\zeta\overline{\pi}\kappa)^{(t-1)s}(1-\overline{\pi})^{1-s}} = (\zeta\overline{\pi}\kappa)^s > 1$ as long as $\zeta\overline{\pi}\kappa > 1$.

Notes

1. Jim Oeppen and James Vaupel (2002, p. 1029) report that 'female life expectancy in the record-holding country has risen for 160 years at a steady pace of almost 3 months per year.'
2. The same is true of discounted utilitarianism, which, as Roemer has privately argued, is particularly unjustified in the intergenerational context. See Geoffrey Heal (1998) for a recent discussion of various normative criteria in dynamic economies.
3. See Roemer (1996, ch. 5). This index should be seen as unrelated to *reported happiness*, a measure which typically fails to respond to improvements to living conditions along time and which displays marked adaptation to unfavourable conditions (see, for example, Daniel McFadden, 2005, and Daniel Kahneman, Ed Diener and Norbert Schwarz, eds, 2003).
4. Veneziani has called my attention to this literature.
5. Veneziani has correctly observed that the idea of applying an homogeneous sufficiency criterion to the seniors of various generations, including the initial seniors, would be unjustified if extreme past exploitation or discrimination demanded special compensation to the initial seniors.
6. Under the natural assumption that the indifference curves of Λ are not flat in a neighbourhood of the vertical axis, where $c_m = 0$, because the path $(c_{s0}, (\beta - c_{s0}, \beta^2), \{(0, \beta^{t+1})|t = 2, \dots\})$ is also feasible, and it gives Generation One, which is the single worst off in path$\{\beta, \{(0, \beta^{t+1})|t = 1, 2, \dots\}$, a higher standard of living than the one obtained in $(\beta, \{(0, \beta^{t+1})|t = 1, 2, \dots\}$, without affecting the standards of living of generations 2, 3, ... obtained there.
7. Alternatively, we could consider that capital depreciates *before* it enters production, so that $K_t = (1 - \delta)\hat{K}_t$. The distinction does not apply to the original Diamond (1965) model, which assumes $\delta = 0$. It also assumes positive population growth.
8. Diamond (1965) assumes an additively separable utility function.
9. But not necessarily for stocks of natural capital: a forest can be cleared and converted into firewood.
10. See Kenneth Arrow (1962), Marvin Frankel (1962), John Chipman (1970), Paul Romer (1986), as well as the extensive references in Philippe Aghion and Peter Howitt (1999).
11. Roemer and Veneziani use Sen's term 'functionings' instead of standard of living.

References

Aghion, Philippe, and Peter Howitt (1999) *Endogenous Growth Theory* (Cambridge, MA: MIT Press).

Anderson, Elizabeth (1999) 'What Is the Point of Equality?', *Ethics*, vol. 109, pp. 287–337.
Arneson, Richard J. (2000) 'Perfectionism and Politics', *Ethics*, vol. 111, pp. 37–63.
Arneson, Richard J. (2002) 'Why Justice Requires Transfers to Offset Income and Wealth Inequalities', *Social Philosophy and Policy*, vol. 19, no. 1, pp. 172–200.
Arneson, Richard J. (2005) 'Distributive Justice and Basic Capability Equality', in Alexander Kaufman (ed.), *Capabilities Equality: Basic Issues and Problems* (London: Routledge).
Arrow, Kenneth J. (1962) 'The Economic Implications of Learning by Doing', *Review of Economic Studies*, vol. 29, no. 1, pp.155–73.
Arrow, Kenneth J., and Anthony C. Fisher (1974) 'Environmental Preservation, Uncertainty and Irreversibility', *Quarterly Journal of Economics*, vol. 88, no. 1, pp. 312–19.
Bishop, Richard C. (1978) 'Endangered Species and Uncertainty: the Economics of a Safe Minimum Standard', *American Journal of Agriculture Economics,* vol. 60, no. 1, pp. 10–18.
Chipman, John S. (1970) 'External Economies of Scale and Competitive Equilibrium', *Quarterly Journal of Economics*, vol. 84, no. 3, pp. 347–85.
Diamond, Peter (1965) 'National Debt in a Neoclassical Growth Model', *American Economic Review*, vol. 55, pp. 1126–50.
Frankel, Marvin (1962) 'The Production Function in Allocation and Growth: a Synthesis', *American Economic Review*, vol. 52, pp. 995–1022.
Frankfurt, Harry (1987) 'Equality as a Moral Idea', *Ethics*, vol. 98, pp. 21–42.
Gersovitz, Mark (1988) 'Savings and Development', in H. Chenery and T. N. Srinivassan (eds), *Handbook of Development Economics,* vol. 1 (Amsterdam: North Holland).
Heal, Geoffrey (1998) *Valuing the Future* (New York: Columbia University Press).
Kahneman, Daniel, Ed Diener, and Norbert Schwarz (eds) (2003) *Well-Being: The Foundations of Hedonic Psychology* (New York: Russell Sage Foundation).
McFadden, Daniel (2005) 'The New Science of Pleasure: Consumer Behaviour and the Measurement of Well-being', Frisch Memorial Lecture, Econometric Society World Congress, London, August 2005.
McKerlie, Dennis (1989) 'Equality and Time', *Ethics,* vol. 99, pp. 475–91.
Nussbaum, Martha C. (2000) 'Aristotle, Politics, and Human Capabilities: A Response to Antony, Arneson, Charlesworth and Mulgan', *Ethics*, vol. 111, pp. 102–40.
Oeppen, Jim, and James W. Vaupel (2002) 'Broken Limits to Life Expectancy', *Science*, vol. 296, pp. 1029–31.
Rawls, John (1971) *A Theory of Justice* (Cambridge, MA: Harvard University Press).
Roemer, John (1996) *Theories of Distributive Justice* (Cambridge, MA: Harvard University Press).
Roemer, John (2004) 'Eclectic Distributional Ethics,' *Politics, Philosophy and Economics*, vol. 3, no. 3, pp. 267–81.
Roemer, John (2007) Ch. 12 in this volume.
Roemer, John, and Roberto Veneziani (2004) 'What We Owe Our Children, They Their Children, …', *Journal of Public Economic Theory,* vol. 6, no. 5, pp. 637–54.
Romer, Paul M. (1986) 'Increasing Returns and Long-run Growth', *Journal of Political Economy*, vol. 94 (5), pp. 1002–37.
Shinotsuka, Tomoichi, Koichi Suga, Kotaro Suzumura and Koichi Tadenuma (2007) Ch. 2 in this volume.

Silvestre, Joaquim (2002) 'Progress and Conservation under Rawls's Maximin Principle', *Social Choice and Welfare*, vol. 19, pp. 1–27.

Steger, Thomas (2002) 'Productive Consumption, the Intertemporal Consumption Trade-off and Growth', *Journal of Economic Dynamics and Control,* vol. 26, pp. 1053–68.

Temkin, Larry (1993) *Inequality* (Oxford: Oxford University Press).

Walzer, Michael (1983) *Spheres of Justice: A Defense of Pluralism and Equality* (New York: Basic Books).

Part V
Long-Run Issues of Intergenerational Equity

15

Toward a Theory of a Just Savings Principle

Ngo Van Long

15.1 Introduction

The question of intergenerational equity has troubled many philosophers, economists and scientists concerned with social issues. Environmentalists have argued that the modern economies are destroying the environment and reducing biodiversity, at the expense of future generations. At the other pole, Immanuel Kant (1724–1804) found it disconcerting that earlier generations should carry the burdens for the benefits of later generations.[1] In a similar vein, John Rawls expressed the opinion that 'the utilitarian doctrine may direct us to demand heavy sacrifices of the poorer generations for the sake of greater advantages for later ones that are far better off. (...) Even if we cannot define a precise just savings principle, we should be able to avoid this sort of extremes.'[2]

In economics, most theorists seem quite comfortable with the utilitarian approach in their investigation of optimal growth and optimal savings for a time horizon extending to infinity, but there is substantial disagreement among them concerning the use of discounting[3]. While it is true that utilitarianism has been the predominant theory that underlies much of modern moral philosophy (and much of modern economics[4]), the principle of utility has been questioned because many of its implications seem at odds with some moral conceptions of justice. The utilitarian objective, namely maximizing the greatest net balance of satisfactions, entails the possibility that the loss of freedom by some is justifiable by a greater good delivered to others. In opposition, Rawls (1958, 1963, 1971) advances the theory of 'Justice as Fairness', which does not permit such tradeoffs. According to justice as fairness, the principles of justices must be principles that free and rational persons 'would accept in an initial position of equality as defining the fundamental terms of their association.'[5]

Rawls' conception of justice has its foundation in the theory of social contract advanced by Locke, Rousseau, and Kant. The initial position conceived by Rawls is a hypothetical situation, in which the contracting parties are individuals hidden behind the veil of ignorance: none of them knows his place in society, his natural talents, intelligence, strength, and the like. In other words, the principles of justice are agreed to in an initial situation that is fair. Rawls argues that the contracting parties would agree to two principles of justice. The first principle says that 'each person is to have an equal right in the most extensive scheme of equal basic liberties compatible with a similar scheme of liberties for others.'[6] The second principle states that social and economic inequalities are acceptable only if they are arranged so that they are 'both (a) to the greatest expected benefit of the least advantaged and (b) attached to offices and positions open to all under conditions of fair equality of opportunity.'[7]

In formulating the second principle, Rawls had in mind the fact that some degree of inequality may have an incentive effect that makes everyone better off. To quote:

> To illustrate . . . consider the distribution of income among social classes . . . Now those starting out as members of the entrepreneurial class . . . have a better prospect than those who begin in the class of unskilled labourers. It seems likely that this will be true even when the social injustices which now exist are removed. What, then, can possibly justify this kind of initial inequality in life prospects? According to the difference principle, it is justifiable only if the difference of expectation is to the advantage of . . . the representative unskilled worker. The inequality in expectation is permissible only if lowering it would make the working class even more worse off . . . The greater expectations allowed to entrepreneurs encourages them to do things which raise the prospects of labouring class. Their better prospects act as incentives so that the economic process is more efficient, innovation proceeds at a faster space, and so on.[8]

Simply put, the second principle is arrived at by combining the principle of fair equality of opportunity with the so-called 'difference principle',[9] which in the two-person case, states that, unless there is a distribution of income that makes both persons better off, an equal distribution is to be preferred. Economists generally refer to Rawls' difference principle as the maximin rule. Many economists have applied this rule to the problem of intergenerational equity. (See, for example, Solow,[10] 1974, Burmeister and Hammond, 1978, Dixit, Hammond and Hoel, 1980, and the references cited therein). Rawls, however, argued emphatically that 'the difference principle does not hold for the questions of justice between generations'.[11] He suggested some considerations (which I review below) that would lay a foundation for a 'just savings principle' which would limit the scope of the difference principle. However, Rawls did not give a formal definition of a just savings principle.

In the present chapter, I review Rawls' objection to the direct application of the difference principle to intergenerational equity, and I propose some adjustments to the difference principle so as to pin down the concept of just savings. Some implications of my proposed adjustments will be examined using some formal economic models. I also compare, in section 15.6, my approach with those of several economists who had proposed some formal models to deal with the questions of Rawlsian just savings, in particular, Arrow (1973), Dasgupta (1974a, 1974b), Leininger (1985), Calvo (1978), Rodriguez (1981), Pezzey (1994), Asheim (1988), Suga and Udagawa (2004, 2005).

15.2 A brief review of Rawls' difference principle

According to Aristotle, justice consists of refraining from (a) gaining some advantage to oneself by seizing what belongs to another, and (b) denying a person what is due to him (for example, the showing of proper respect). This is basically derived from Plato's idea of justice, according to which a just person does not overstep the boundary of his sphere. As Rawls points out, Aristotle's definition 'clearly presupposes . . . an account of what properly belongs to a person and of what is due to him.'[12] It follows that a theory of social justice must be concerned with the principles of 'assigning rights and duties' and 'defining appropriate division of social advantages'.[13]

The ancient Greeks did not equate justice with equality (neither *ex ante* equality, nor *ex post* equality): it was thought just that slaves were slaves. The modern utilitarian view of justice displays *ex ante* equality: the sum of individual utilities must be maximized; every term in the sum receives the same weight.[14] *Ex post*, however, some individuals would be forced to suffer for the 'greater good': torture of a suspect may be justified on utilitarian grounds. This view offends the moral sense of many. In opposing the utilitarians, Rawls proposes the contractual approach in which justice is based on fairness of a hypothetical original position, and argues that the contracting parties, hidden behind the veil of ignorance, will choose his two principles of justice.

There are two separate issues concerning the contractual approach to a theory of justice. The first issue is about the interpretation of the hypothetical original situation (e.g. who are the contracting parties, and what is their problem of choice). The second issue is what the parties would choose as principles of justice.[15] The principles are just if they appeal to our sense of justice, which is arrived at in a 'reflective equilibrium'.[16] When one abstracts from intergenerational considerations, one can posit that the persons in the hypothetical original situation are rational[17] individuals who care only about themselves (this is the assumption of mutually disinterested rationality), but none of them knows 'his place in society, his class position or social status; nor does he know his fortune in the distribution of natural assets and abilities.'[18] They assume that they 'normally prefer more primary social goods rather than less.'[19] According to Rawls, the original position is defined so that we get the 'desired solution.'[20]

When it comes to justice between generations, Rawls suggests that we must modify our assumptions concerning the original position, by specifying that the parties in the original position are 'heads of families and therefore have a desire to further the well-being of at least their more immediate descendants.'[21] As I argue below, such a formulation would introduce some elements of utilitarianism into the Rawlsian framework. Rawls also introduces the constraint that the just savings principle adopted must be such that the parties wish all earlier generations to have followed it.[22]

Why is it necessary to bring such modifications to the difference principle? The answer is simple: the unmodified difference principle would entail 'either no saving at all or not enough saving to improve social circumstances'.[23] Such a state of affairs would offend our sense of justice (or at least, Rawls' sense of justice). In other words, one must modify assumptions so as to 'achieve a reasonable result'.[24] It is not clear how to judge if a result is 'reasonable.' It seems that one must rely on intuition, even though Rawls remarks that 'an intuitionist conception of justice is, one might say, but half a conception.'[25] Rawls suggests that a just savings principle should limit the scope of the difference principle (p. 258). He acknowledges the impossibility of being very specific about the range of savings rate, but hopes that certain extremes will be excluded: 'Thus we may assume that the parties avoid imposing very high rates [of saving] at the earlier stages of accumulation.'[26]

In what follows, I propose some modifications of the difference principle to deal with intergenerational equity. Since I wish to focus on justice between generations, I abstract from intragenerational justice, by assuming that within each generation, all individuals are identical and are treated identically (both *ex ante* and *ex post*). I suggest two alternative modifications to the difference principle when applied to justice between generations. In the first alternative, I posit that, since individuals care about their immediate descendants, the welfare of an individual consists not just of his utility of his consumption of the primary goods, but also on the extent to which his son's consumption surpasses his.[27] Thus I make the distinction between a person's utility of consumption and his welfare. Applying the difference principle to welfare levels rather than utility-of-consumption levels, I show that the maximin objective is consistent with positive net savings and an increasing time path of consumption. I call this formulation 'Just Savings Principle with Care for Immediate Descendants.' In this formulation, one needs not assume that the father knows the preferences of his son.

My second alternative is called 'A Dynastic Approach with Concern for the Least Advantaged'. In this formulation, there is a tradeoff between the utility stream of a dynasty (a family line extending to the indefinite future) and the utility level of the least advantaged member of that dynasty. I shall show that this formulation does in fact ensure that the parties avoid imposing very high rates at the earlier stages of accumulation. Before investigating in detail the implications of these two approaches, it is instructive to study the consequences of applying the unmodified difference principle to intergenerational equity.

15.3 Implications of the unmodified difference principle for savings and growth

Rawls remarked that if we do not modify the assumptions about the original positions, the contracting parties would not want to undertake any saving. By assumption, the persons in the original position know that they are contemporaries,[28] even though they do not know the particular circumstances of their society. Being selfish (by assumption) they will agree that 'no one has a duty to save for posterity'.[29] He then stated that 'so in this instance the veil of ignorance fails to secure the desired result'.[30] It is clear that by 'the desired result,' Rawls meant some saving should take place, at least when the economy is still at a low level of development. I shall set up a few models to show that the application of the (unmodified) difference principle does in fact imply zero saving in the sense that net capital formation is zero. While it is clear that 'capital is not only factories and machines, and so on, but also knowledge and culture',[31] for the sake of simplicity, I begin with a model with a single capital stock.

Model 1: A single capital stock

Consider a continuous time model. Each individual lives for just one instant. Within each generation, all individuals are identical and are treated identically. The population is constant. The representative individual of generation t consumes $c(t)$ and obtains the utility level $u(c(t))$. Here $u(.)$ is an increasing function. The capital stock $k(t)$ depreciates at the rate $\delta > 0$. The output is

$$y(t) = f(k(t))$$

where $f(.)$ is a strictly concave production function, with $f(0) = 0$, $f'(0) > \delta$ and $f'(\infty) < \delta$. The rate of growth of the capital stock is

$$\dot{k}(t) = f(k(t)) - c(t) - \delta k(t)$$

where $k(0) = k_0 > 0$. Feasibility requires that $k(t) \geq 0$ for all t.

The (unmodified) difference principle requires the maximization of the utility of consumption of the least advantaged generation.

Definition 1: A non-negative time path $c(.)$ is a feasible consumption path if the differential equation

$$\dot{k}(t) = f(k(t)) - c(t) - \delta k(t)$$

has a solution with $k(0) = k_0$ and $k(t) \geq 0$ for all t.

Definition 2: c_m is said to be the worst consumption level along the feasible consumption path $c(.)$ if

$$c_m = \inf\{c(t); t \in [0, \infty)\}$$

Definition 3: The maximin performance of the consumption path $c(.)$ is

$$P(c(.)) = u(c_m)$$

Let us denote by $S(k_0)$ the set of all feasible consumption paths (given the initial stock k_0).

Definition 4: A feasible consumption path $c^*(.)$ in $S(k_0)$ is said to satisfy the difference principle if and only if

$$P(c^*(.)) \geq P(c(.))$$

for all $c(.) \in S(k_0)$.

The following proposition can then be proved:

Proposition 1: For all $k_0 \in (0, \widehat{k}_G)$ where $\widehat{k}_G$ satisfies $f'(\widehat{k}_G) = \delta$, the only feasible consumption path that satisfies the difference principle is the constant consumption path obtained by setting net capital formation at $\dot{k} = 0$ for ever:

$$c(t) = f(k_0) - \delta k_0 \text{ for all } t \in [0, \infty)$$

Proof: Any alternative path $\widetilde{c}(.)$ with $\widetilde{c}(0) < f(k_0) - \delta k_0$ is inferior to the constant path described above, because the minimum consumption along such a path is lower than $f(k_0) - \delta k_0$. Any alternative path $\widetilde{c}(.)$ with $\widetilde{c}(0) > f(k_0) - \delta k_0$ will result in a fall in the capital stock below k_0, which, since $k_0 < k_G$, entails $\widetilde{c}(t) < f(k_0) - \delta k_0$ for some $t > 0$, in view of the feasibility requirement that $k(t) \geq 0$ for all t.

Remark 1: Proposition 1 makes precise the Rawlsian assertion that the unmodified difference principle would lead to 'no saving at all'. The absence of saving is a concern for Rawls, especially if one is considering a society with a very low initial level of capital. Why? Because, 'to establish effective just institutions within which the basic liberties can be realized',[32] a society must have a sufficient material base. Generations must 'carry their fair share of the burden of realizing and preserving a just society'.[33]

Remark 2: Rawls does not say that accumulation should go on for ever. He has in mind the eventual attainment of a long-run stationary state: 'Eventually, once just institutions are firmly established and all the basic liberties effectively realized, the net accumulation asked for falls to zero'.[34]One must not think that such a Rawlsian long-run stationary state corresponds to the 'golden rule' capital stock level $\hat{k}$. As Rawls puts it, the last stage is 'not one of great abundance... It is a mistake to believe that a just and good society must wait upon a high standard of life. What men want is meaningful work in free association with others, these associations regulating their relations to one another within a framework of just basic institutions. To achieve this state of things great wealth is not necessary. In fact, beyond some point it is more likely to be a positive hindrance, a meaningless distraction at best if not a temptation to indulgence and emptiness.'[35]

Model 2: a multi-capital-stock model

In the preceding subsection, there is only one capital stock, so the concept of 'zero net capital formation' has no ambiguities. If we turn to a model with several distinct capital stocks, zero net capital formation should be understood in the value sense: a capital stock (say, fossil oil) may be allowed to fall, as long as there are sufficient accumulations of other capital stocks (e.g. solar panels) to compensate for it, so that in value terms, net accumulation is zero. How does one value a capital stock? In what follows, I take it that the value of a capital stock is the 'correct' price (per unit) multiplied by the number of units. I have used the qualification 'correct', because, as is generally known, there are many instances where the market price is not the correct price. For example, the correct price of a forest would have to reflect not only its timber value but also its contribution to biodiversity etc. Without correct pricing, an economist would be, in the words of Oscar Wilde, 'someone who knows the price of everything and the value of nothing.'[36]

I do not discuss further the issue of correct pricing. In what follows, I simply assume that any market failure has been corrected by appropriate measures; see Asheim (2000). I now introduce a model with two capital stocks and one consumption good. Generalization to any number $n > 2$ of stocks is straightforward. Suppose there are two stocks of capital: a man-made capital denoted by K, and a stock of natural capital, denoted by X. The net investments in these stocks are denoted by $\dot{K} = K$ and $\dot{X}$. The economy's transformation surface is defined by

$$C = T(K, \dot{K}, X, \dot{X}) \tag{15.1}$$

where C is the output of the consumption good. As usual, assume that $T_{\dot{K}} < 0$ and $T_{\dot{X}} < 0$ because more current investments imply less current consumption.

Define the (implicit) price of a stock to be the cost of investment in that stock in terms of forgone consumption:

$$P_K \equiv -T_{\dot{K}} > 0$$

$$P_X \equiv -T_{\dot{X}} > 0$$

Define the (implicit) rental rate of a stock to be marginal contribution of that stock to the flow of consumption good:

$$R_K \equiv T_K$$
$$R_X \equiv T_X$$

Definition 5: The net rate of return of an asset i (denoted by ρ_i) is the sum of (a) the proportional rate of change in its asset price (i.e., the rate of capital gain) and (b) the rental rate per dollar invested in the asset:

$$\rho_i \equiv \frac{\dot{P}_i}{P_i} + \frac{R_i}{P_i} \quad \text{where } i = K, X$$

Definition 6: The no-arbitrage condition is said to hold along a path if the net rates of return are the same for all assets:

$$\frac{\dot{P}_K}{P_K} + \frac{R_K}{P_K} = \frac{\dot{P}_X}{P_X} + \frac{R_X}{P_X} \equiv \rho(t) \tag{15.2}$$

Remark 3: The no-arbitrage condition may also be written as:

$$\frac{1}{T_{\dot{K}}}\frac{dT_{\dot{K}}}{dt} - \frac{T_K}{T_{\dot{K}}} = \frac{1}{T_{\dot{X}}}\frac{dT_{\dot{X}}}{dt} - \frac{T_X}{T_{\dot{X}}} \tag{15.3}$$

Remark 4: If the no-arbitrage condition holds, the economy is said to be dynamically efficient between periods. Overall dynamic efficiency requires, in addition to the no-arbitrage condition, a transversality condition, which prevents over-accumulation or explosive deficits.

In this section, I consider the following transversality condition

$$\lim_{t\to\infty} N(t)\exp\left[-\int_0^t \rho(\tau)d\tau\right] = 0 \tag{15.4}$$

Roughly speaking, a feasible path (C^*, K^*, X^*) is said to satisfy dynamic efficiency between periods if, for any finite time interval $[t_1, t_2)$ where $0 \le t_1 < t_2 < \infty$, and for fixed $(K^*(t_1), X^*(t_1), K^*(t_2), X^*(t_2))$, it is not possible to find a feasible path (C, K, X) that both satisfies the condition $(K(t_1), X(t_1), K(t_2), X(t_2)) = (K^*(t_1), X^*(t_1), K^*(t_2), X^*(t_2))$ and achieves higher consumption over any time interval $[\tau_a, \tau_b) \in [t_1, t_2)$ without reducing consumption over some other time interval $[\tau_c, \tau_d) \in [t_1, t_2)$. A feasible path (C^*, K^*, X^*) is said to satisfy overall dynamic efficiency if for starting stocks$(K^*(0), X^*(0))$, it is not possible find a feasible path (C, K, X) that both satisfies the condition $(K(0), X(0)) = (K^*(0), X^*(0))$ and achieves higher consumption over any time interval $[\tau_a, \tau_b) \in [t_1, t_2)$ without reducing consumption over some other time interval $[\tau_c, \tau_d) \in [t_1, t_2)$. Take the case where there is no depreciation and marginal product is positive. Then the strategy of investing all output in the stock K, such that $C(t) = 0$ for all $t \geq 0$ yields a path that is dynamic efficient between periods, but that fails to achieve overall dynamic efficiency. Overall dynamic efficiency is a necessary condition for Pareto efficiency (but may not be sufficient, because the latter concept may require a specification of utility function if individuals live over finite periods and have preferences over the shape of their consumption paths.)

Definition 7: Net investment, denoted by N is the value sum of changes in all stocks:

$$N = P_K\dot{K} + P_X\dot{X} = -T_{\dot{K}}\dot{K} - T_{\dot{X}}\dot{X} \tag{15.5}$$

Proposition 2: Assume that the no-arbitrage condition holds along a feasible path of consumption and investment. If along that path net investment is zero for ever (i.e. $N(t) = 0$ for all $t \geq 0$), then consumption is constant over time.

Proof: The rate of change in net investment is

$$\dot{N} = -\dot{K}\frac{dT_{\dot{K}}}{dt} - T_{\dot{K}}\frac{d\dot{K}}{dt} - \dot{X}\frac{dT_{\dot{X}}}{dt} - T_{\dot{X}}\frac{d\dot{X}}{dt} \tag{15.6}$$

The rate of change in consumption is, from (15.1)

$$\dot{C} = T_K\dot{K} + T_{\dot{K}}\frac{d\dot{K}}{dt} + T_X\dot{X} + T_{\dot{X}}\frac{d\dot{X}}{dt} \tag{15.7}$$

Adding (15.7) to (15.6), and using the no-arbitrage condition (15.3), we obtain

$$\begin{aligned}\dot{C} + \dot{N} &= -\dot{K}T_{\dot{K}}\left[\frac{1}{T_{\dot{K}}}\frac{dT_{\dot{K}}}{dt} - \frac{T_K}{T_{\dot{K}}}\right] - \dot{X}T_{\dot{X}}\left[\frac{1}{T_{\dot{X}}}\frac{dT_{\dot{X}}}{dt} - \frac{T_X}{T_{\dot{X}}}\right] \\ &= N\left[\frac{1}{T_{\dot{K}}}\frac{dT_{\dot{K}}}{dt} - \frac{T_K}{T_{\dot{K}}}\right] = N\rho \end{aligned} \tag{15.8}$$

From equation (15.8) $N(t) = 0$ for all t implies $\dot{C}(t) = 0$ for all t.

Remark 5: Proposition 2, which is a version of Hartwick's rule, can be proved by other methods, see, for example, Dixit, Hammond and Hoel (1980). For a critical discussion of the implications of this result, see Asheim, Buchholtz, and Withagen (2003).

Proposition 3: Assume that the no-arbitrage condition and the transversality condition (15.4) hold along a feasible path of consumption and investment. If along that path consumption is constant for ever, then net investment is zero for ever.

Proof: From equation (15.8), if $\dot{c}(t) = 0$ for all $t \geq 0$ then

$$N(t) = N_0 \exp\left[\int_0^t \rho(\tau)d\tau\right] \tag{15.9}$$

hence

$$\lim_{t\to\infty} N(t)\exp\left[-\int_0^t \rho(\tau)d\tau\right] = N_0 \tag{15.10}$$

But the transversality condition (15.4) requires

$$\lim_{t\to\infty} N(t)\exp\left[-\int_0^t \rho(\tau)d\tau\right] = 0 \tag{15.11}$$

It follows that $\dot{C} = 0$ for all $t \geq 0$ implies $N_0 = 0$, and hence $N(t) = 0$ for all $t \geq 0$.

Remark 6: A somewhat different version of Proposition 3 was provided by Withagen and Asheim (1998). Dixit, Hammond and Hoel (1980) use a different set of sufficient conditions.

Example 1:
This example shows how the transformation function surface $C = T(K, \dot{K}, X, \dot{X})$ can be derived from more basic production relations, and how net saving can be computed. In this example, K is a man-made capital stock, and the stock X represents the quality of the environment. Assume the final good is produced using three inputs: K, X, and E, where $E \geq 0$ is the emission of pollutants. Let Y denote the output of the final good. The production function is

$$Y = F(K, E, X)$$

with

$$F_K > 0$$
$$F_X \geq 0$$
$$F_E > 0$$

Assume X evolves according to the dynamic law

$$\dot{X} = G(X) - E$$

where E causes the deterioration of X that arises from the production of the final good, and $G(X)$ is a function describing the natural rate of regeneration of the environment. Output is allocated between consumption and investment in man-made capital

$$C + I = F(K, E, X)$$

There is no depreciation of man-made capital, so that

$$\dot{K} = I$$

The transformation surface $C = T(K, \dot{K}, X, \dot{X})$ is derived from the basic features of the economy as follows:

$$C = F(K, E, X) - \dot{K} = F(K, G(X) - \dot{X}, X) - \dot{K} \equiv T(K, \dot{K}, X, \dot{X}) \tag{15.12}$$

Thus, in this model,

$$\begin{aligned}
R_K &= T_K = F_K > 0 \\
P_K &= -T_{\dot{K}} = 1 \\
R_X &= T_X = F_E G'(X) + F_X \\
T_{\dot{X}} &= -F_E < 0 \\
\rho(t) &= F_K(K(t), E(t), X(t)).
\end{aligned}$$

It follows that net saving is

$$N \equiv \dot{K} + F_E \dot{X}$$

With our special form of T as given by (15.12), we obtain, using (15.7),

$$\dot{C} + \frac{d}{dt}[\dot{K} + F_E \dot{X}] = [\dot{K} + F_E \dot{X}] F_K$$

hence

$$\dot{K} + F_E \dot{X} = \frac{\dot{C}}{F_K} + \frac{1}{F_K} \frac{d}{dt}[N] \tag{15.13}$$

It is easy to verify that, under non-arbitrage, if net investment $N(t)$ is zero for ever, then consumption is constant for ever. Conversely, if consumption is constant for ever, then $\dot{N} = N F_K$, which, together with the transversality condition, implies $N(t) = 0$ for all t.

15.4 A modified difference principle: care for immediate descendants

I have shown in the preceding section that the unmodified difference principle leads to zero saving, no matter how small is the initial capital stock. Without capital accumulation, there is little chance that just institutions can be developed and sustained. To obtain reasonable results, the difference principle must be modified: it must be supplemented by a just savings principle. Rawls acknowledges that it is difficult to formulate a just savings principle. At the same time, certain reasonable assumptions would set limits on the savings rate. Thus, in dealing with intergenerational equity, Rawls assumes that the parties in the original situation are heads of family, and that the principle adopted must be such that the parties wish all earlier generations to have followed it. 'Thus imagining themselves to be fathers, say, they are to ascertain how much they should set aside for their sons and grandsons by noting that they would believe themselves entitled to claim of their fathers and grandfathers'.[37]

A continuous time model

In this section, I attempt to capture this idea in a simple continuous time model. Each individual lives for just one instant. The representative individual of generation t consumes $C(t)$. Take $\dot{C}(t)$ as the rate of change in consumption across adjacent generations. A positive $\dot{C}(t)$ signifies that the son's consumption exceeds that of the father. I posit that the welfare of the representative individual of generation t is an increasing function of both $C(t)$ and $\dot{C}(t)$.

$$u(t) = u(C(t), \dot{C}(t))$$

Here, u denotes a person's welfare, rather than the utility of his consumption. This formulation implies that a father would be willing to accept a small reduction in his consumption if that would result in a sufficiently large increase his son's consumption. I shall refer to $\dot{C}(t)$ as the 'social progress' at time t.

It seems reasonable that the parties would agree that savings should be such that the welfare of the least advantaged generation is maximized. Then the mathematical programme facing the society can be formulated as follows. Choose the highest number $\overline{U}$ such that $u(C(t), \dot{C}(t)) \geq \overline{U}$ for all $t \in [0, \infty)$, where $C(t)$ belongs to the set of feasible consumption paths. Formally, the objective of the social planner is

$$\max \overline{U}$$

such that

$$u(C(t), \dot{C}(t)) \geq \overline{U} \text{ for all } t \in [0, \infty)$$

and subject to the resource constraint.

Analysis of an example

Example 2:
Assume the technology and endowments are the same as in Example 1. Assume the welfare function for the representative individual of generation t is

$$u(C,\dot{C}) = u(C+\alpha\dot{C}) \equiv U(Q)$$

where $Q = C+\alpha\dot{C}$, where $\alpha > 0$ is the marginal rate of substitution between C and $\dot{C}$.

Define

$$z = \dot{C}$$

The modified difference principle requires the solution of the following problem

$$\max \overline{U}$$

subject to

$$u(C,z) \geq \overline{U}$$

Net saving in terms of the consumption good is

$$N(t) = \dot{K}(t) + (F_E)\dot{X}(t)$$

What can we say about the behaviour of net saving along a constant welfare path?

The following proposition can be proved:

Proposition 4: Under the assumptions of Example 2, along any efficient constant welfare path with rising consumption, it is necessary that net investment $N(t)$ be positive and satisfy the condition

$$N(t) = \dot{K}(t) + F_E\left[G(X(t)) - E(t)\right] = \dot{C}(t)\gamma(t)$$

where $\gamma(t)$ is given by:

$$\gamma(t) = \int_t^{\infty} e^{-\int_t^{\tau} \frac{1}{\alpha}+F_K(s)ds} d\tau > 0$$

In particular, if the production function is Cobb-Douglas and is independent of X, and $G(X) = 0$, then starting with $X_0 > 0$, along the maximin welfare path, consumption is rising ($\dot{C} > 0$). But $\dot{C}$ tends to zero as t tends to infinity.

Proof: See the Appendix.

Some open questions concerning the variable representing social progress

In the continuous time model, I have interpreted the expression $\dot{C}(t)$ as 'social progress' . A question that naturally arises is the meaning of this term if one thinks of continuous time as a limiting case of discrete time. If the model were set in discrete time, one can make a distinction between 'looking backward' and 'looking forward.' We may define 'forward looking progress' as

$$\frac{C(t+h)-C(t)}{h}, \; h > 0$$

and 'backward looking progress' as

$$\frac{C(t)-C(t-h)}{h}, \; h > 0.$$

The utility function of the representative individual of generation t could be written as

$$U\left(C(t), \frac{C(t+h)-C(t)}{h}, \frac{C(t)-C(t-h)}{h}\right)$$

where $h > 0$. By taking the limit we have

$$\lim_{h\to 0^+}\left[\frac{C(t+h)-C(t)}{h}\right] = \dot{C}(t^+)$$

$$\lim_{h\to 0^+}\left[\frac{C(t)-C(t-h)}{h}\right] = \dot{C}(t^-)$$

In principle, one can distinguish $\dot{C}(t^+)$ from $\dot{C}(t^-)$. An individual may be want to take actions to increase his son's consumption, and derive pleasure from a greater $\dot{C}(t^+)$. He may not care at all about $\dot{C}(t^-)$. However, for optimal problems that satisfy certain regularity conditions, these two magnitudes are the same almost everywhere. This is clearly a shortcoming of the continuous time formulation. There is a related question that I have not explored. In a discrete time formulation, what would a maximin path look like if the utility function is increasing in both forward looking progress and backward looking progress? Arrow (1973) has shown that if utility of the father is of the following additive separable form

$$U(C_t, C_{t+1}) = v(C_t) + \beta v(C_{t+1})$$

where $v(.)$ is strictly concave, then the maximin path can display the saw-tooth pattern. Does the extension of utility to

$$U(C_t, C_{t-1}, C_{t+1}) = v(C_t) + \beta v(C_{t+1}) + \gamma v(C_{t-1})$$

greatly affect the Arrow's result about the time path of consumption when the social planner pursues the maximin objective? And, if one adopts the more general form (instead of the additive separable form), and take the limit to continuous time, does the optimal time path in the discrete time model approach that of the continuous model?

15.5 An alternative just savings principle: a dynastic approach with concern for the least advantaged

In the preceding section, I argued that a savings rule may be deemed just if it results in the maximization of the welfare of the least advantaged generation, where a person's welfare depends on both his consumption and his perception of 'social progress' as captured by $\dot{C}$. That objective seems reasonable if we assume that the parties are fathers who care for their immediate descendants.

The mixed Rawlsian–Utilitarian objective

Consider now the alternative assumption that the contracting parties are family lines. It is arguable that each party would take into account two things: (i) the consumption level of the least advantaged generation, and (ii) the sum of weighted utility across all generations. It seems also sensible to allow a trade-off between (i) and (ii) above, because each party now represents a family line. The standard utilitarian tradition would treat a family line as an infinitely-lived individual. This could, however, result in requiring great sacrifices of early generations who are typically poor. In contrast, the approach proposed here avoids imposing very high rates (of savings) at the earlier stages of accumulation.

The objective is, for a given λ in [0, 1], to maximize the weighted average of (i) the utility of the least advantaged generation, and (ii) the life-time utility of the fictitious infinitely-lived individual with a given discount rate $\rho > 0$. Thus the social welfare function is:

$$W = \max \lambda \overline{U} + (1 - \lambda) \int_0^\infty e^{-\rho t} U(C(t)) dt \qquad (15.14)$$

The maximization is performed by choosing a number $\overline{U}$ and a consumption and investment path that satisfies the constraint

$$U(C(t)) \geq \overline{U} \text{ for all } t \geq 0$$

and the technology and resource constraints. Since the objective function is a weighted average between a conventional utilitarian term (an integral of discounted utility flow) and a Rawlsian term ($\overline{U}$) that expresses the concern for the least advantaged, we shall call this objective function 'the mixed Rawlsian–Utilitarian objective'.

Clearly, this welfare function can be seen as a compromise between the maximin criterion (which can be obtained by setting $\lambda = 1$ and thus $1 - \lambda = 0$) and the standard utilitarian criterion with discounting (which can be obtained by setting $\lambda = 0$ and thus $1 - \lambda = 1$). However, the reason for proposing this criterion is not a compromise just for the sake of compromising. Rather, the virtue of this new criterion is that it reflects the dual nature of a family line. A family line is at the same time 'one' and 'many'. Being 'one' , it is like a single individual. There are no valid reasons to object to an individual's discounting of his future consumption. This justifies the second part of the welfare function $W(.)$. But a family line is also 'many'. As such, the worse-off individuals have special claims not unlike the those accorded to the 'contemporaneous individuals' of the simple Rawlsian model without intergenerational considerations.

Remark 7:
There remains an open question: is there a reasonable set of axioms that would imply the mixed Rawlsian–Utilitarian objective? A careful study the axiomatization of the mixed Long-Run Average and Discounted-Utilitarian objective by Chichilnisky (1996), as well as the axiomatization of the maximin objective by Strasnik (1976) might be a useful preliminary step toward an answer to that question.

An example of optimal savings under the mixed Rawlsian–Utilitarian objective

In what follows, I present a simple model of economic growth, and show that the use of the welfare criterion $W(.)$ does generate consumption/investment paths that seem quite appealing to our notion of justice.

Example 3: Assume an economy with a single capital stock, denoted by K. The production function is $F(K)$ with positive and diminishing marginal product of capital. Let $\delta > 0$ be the rate of depreciation. Gross investment is $I = F(K) - C$. Then a feasible consumption path $C(.) > 0$ must satisfy

$$\dot{K} = F(K) - C - \delta K$$
$$K(t) \geq 0,\ K(0) = K_0 > 0$$

Our main result for this model can now be stated:

Proposition 5: Under the assumptions of Example 3, the solution of the planning problem under the mixed Rawlsian–Utilitarian objective has the following properties:

(A) If K_0 is below the modified-golden-rule capital stock $\widehat{K}$ (where $F'(\widehat{K}) = \rho + \delta$), the solution will require positive savings, but for those generations at the earlier stages of accumulation, the required savings are less than what the standard utilitarian objective would ask for.

Consumption will be constant over some initial time interval, but eventually consumption will rise, and the capital stock will approach the modified-golden-rule level $\widehat{K}$.

(B) If the economy starts with an initial K_0 that is greater than the modified-golden-rule level $\widehat{K}$, the solution will be to approach monotonically an intermediate capital stock $\overline{K}$ where $K_0 > \overline{K} > \widehat{K}$. Once $\overline{K}$ has been reached, consumption will remain constant.

Proof:
Let us look at the necessary conditions. Following Leonard and Long (1992, Chapter 9), re-write the objective function (15.14) as follows:

$$\max \int_0^\infty [\rho\lambda\overline{U} + (1-\lambda)U(C)]e^{-\rho t}dt$$

Let ψ be the current-value co-state variable associated with the state variable K. The current-value Hamiltonian is

$$H = \rho\lambda\overline{U} + (1-\lambda)U(C) + \psi[F(K) - C - \delta K]$$

Let μ be the multiplier associated with the constraint $U(C(t)) \geq \overline{U}$. The Lagrangian is

$$L = H + \mu[U(C) - \overline{U}]$$

The necessary conditions are

$$\frac{\partial L}{\partial C} = (1-\lambda+\mu)U'(C) - \psi(t) = 0$$

$$\mu \geq 0,\ U(C) - \overline{U} \geq 0,\ \mu[U(C) - \overline{U}] = 0$$

$$\dot{\psi} = [\rho + \delta - F'(K)]\psi$$

$$\dot{K} = F(K) - C - \delta K$$

$$\int_0^\infty e^{-\rho t}\frac{\partial L}{\partial \overline{U}}dt = 0$$

The last equation implies

$$\int_0^\infty [\rho\lambda - \mu(t)]e^{-\rho t}dt = 0 \qquad (15.15)$$

i.e.

$$\lambda = \int_0^\infty \mu(t)e^{-\rho t}dt \tag{15.16}$$

Since $\lambda > 0$, it follows from (15.15) that $\mu(t) > 0$ over some time interval $[t_a, t_b]$. When $\mu(t) > 0$, we have $C = \overline{C}$ where $U(\overline{C}) = \overline{U}$.

To find a solution, consider two mutually exclusive cases. In Case A, we suppose $K_0 < \widehat{K}$. In Case B, we have $K_0 > \widehat{K}$. Please see Appendix II for detail.

15.6 Related literature on Rawlsian just saving

Our objective function in section 15.4 (care for immediate descendants) is related to, but different from the literature that sprang from the so-called Arrow–Dasgupta model (see Arrow, 1973, Dasgupta, 1974b). Since the unmodified difference principle would imply zero saving, Arrow and Dasgupta proposed a model where each generation's welfare is an additively separable function of its own consumption and consumption of the next generation:

$$W_t(C_t, C_{t+1}) = U(C_t) + \beta U(C_{t+1})$$

where $\beta \leq 1$ and $U(.)$ is strictly concave. They assumed that technology is linear

$$K_{t+1} = \alpha(K_t - C_t) \quad \text{where } \alpha > 1$$

The objective function of the social planner in the Arrow–Dasgupta model is to find a feasible infinite sequence $(C_0, C_1, \ldots, C_t, \ldots)$ and a largest number $\overline{W}$ such that

$$W_t(C_t, C_{t+1}) \geq \overline{W} \quad \text{for all } t \in \{0, 1, 2, \ldots\}$$

This formulation allows growth to take place.[38] The Arrow–Dasgupta model has the property that the solution of the optimization problem is generally time-inconsistent: if the planner can replan at some time $t_1 > 0$, he would not choose the same sequence, see Dasgupta (1974b), Calvo (1978), Leininger (1985). This type of time-inconsistency arises even in more general models, where W_t is non-decreasing in each C_{t+j} where $j = 1, 2, \ldots, n$. In contrast, our welfare function in section 15.4 is $W_t = u(C(t), \dot{C}(t))$ where both partial derivatives of u are positive, or, in discrete time, $W_t = u(C_t, C_{t+1} - C_t)$. Thus, in the discrete time version, for a given C_{t+1}, W_t may be non-monotone in C_t. It is not clear if time-inconsistency would arise in our model.

One way of avoiding time-inconsistency (at least in the context of the familiar one-sector neoclassical technology) is to use non-paternalistic altruism: the welfare of each individual is the sum of a function of current consumption and the welfare (rather than utility of consumption) of the next generation. See Calvo (1978) and Rodriguez (1981). However, Asheim

(1988) shows that if technology includes both man-made capital and a non-renewable resource, then time-inconsistency reappears. This motivates setting up a game-theoretic model. In such a model (see, for example, Lane and Mitra, 1981, and Asheim, 1988), the current generation chooses a strategy which is best reply of the strategies of future generations. Asheim's main result is that in this game-theoretic setting, the equilibrium programmes have the property of maximizing altruistic utility over the class of feasible programmes with non-decreasing consumption. In Asheim's paper, generations retain their commitment to maximin as an ethical principle. This is in sharp contrast to the usual game-theoretic models where authors[39] assume that each generation maximizes its altruistic utility.

The objective function we consider in section 15.5 is 'the mixed Rawlsian–Utilitarian objective'. This has some similarity with the objective function proposed by Chichilnisky (1996) for sustainable development: she wants to maximize a weighted average of two terms: a conventional utilitarian term (integral of discounted utility flow) and a term that reflects only the long-run property of a consumption sequence. A major problem with her proposed objective function is that an optimal path may fail to exist even in environments where existence would be ensured if the weights are unity for one term, and zero for the other term (see Beltratti, Chichilnisky, and Heal, 1994). Note that Chichilnisky's objective function is derived from a set of axioms. It seems that our mixed Rawlsian–Utilitarian objective can be axiomatized in a similar fashion.[40]

15.7 Concluding remarks

In this chapter, taking seriously the Rawlsian considerations for intergenerational equity, I have proposed two approaches that would yield savings rules that could be deemed just. These approaches give rise to different objective functions. I have provided examples and characterized the solution under the derived objective functions. It is worth noting that while Rawls vehemently opposes the utilitarian principle, his desire to avoid imposing very high rates of savings at the earlier stages of accumulation could not be accommodated without admitting in, perhaps via the back door, some weak form of utilitarian consideration. Such admission of utilitarian elements is implicit in my first approach, and more explicit in my second approach.

An important issue that I have not explored is the possibility of continuous replanning. If this is allowed, the question of time-consistency would have to be faced. Under what conditions would the maximin welfare level $\overline{U}$ be changing over time? If it is anticipated to change over time, how would the current generation take this into account, given the objective functions that I have proposed? This is an interesting topic for future research.

Another interesting issue is how the just savings principle must be adjusted in a globalized world where countries share common environmental assets.

Analysis of such issues would require techniques of differential games, see Dockner *et al.* (2000).

Acknowledgements

I would like to thank Geir Asheim, Marc Fleurbaey, Reiko Gotoh, Kim Long, Tapan Mitra, Wlodek Rabinowicz, John E. Roemer, Koichi Suga, Kotaro Suzumura, Koichi Tadenuma, and Peter Vallentyne for comments on an earlier version of this chapter. Support from the Social Sciences and Humanities Research Council and the Fonds qúébécois de la recherche sur la société et la culture are gratefully acknowledged.

Appendix I: Proof of Proposition 4

Note that for any positive number ρ, the following equality holds:

$$\int_0^\infty \rho \overline{U} e^{-\rho t} dt = \overline{U}$$

Thus, following Leonard and Long (1992, Chapter 9), our problem is equivalent to finding

$$V \equiv \max_{\overline{U}} \int_0^\infty \rho \overline{U} e^{-\rho t} dt \tag{15.17}$$

subject to

$$u(C, z) - \overline{U} \geq 0$$

$$\dot{K} = F(K, E) - C,\ K(0) = K_0$$

$$\dot{X} = G(X) - E,\ X(0) = X_0$$

$$\dot{C} = z$$

and $K(t) \geq 0$, $X(t) \geq 0$, $C(t) \geq 0$.

Note that C is treated as a state variable. Let π_K, π_X, and π_C be the co-state variables associated with the state variables K, X, and C. The Hamiltonian is

$$H = \rho \overline{U} + \pi_K[F(K, E) - C] + \pi_X[G(X) - E] + \pi_C[z] \tag{15.18}$$

and the Lagrangian is

$$L = H + \omega[u(C, z) - \overline{U}]$$

where $\omega \geq 0$ is the multiplier associated with the constraint $u(C, z) - \overline{U} \geq 0$.

The optimality conditions are

$$
\begin{aligned}
&\frac{\partial L}{\partial E} = \pi_K F_E - \pi_X = 0 \\
&\frac{\partial L}{\partial z} = \pi_C + \alpha\omega u_z(C, z) = 0 \text{ which implies } \pi_C \le 0 \\
&\omega \ge 0,\, u(C, z) - \overline{U} \ge 0,\, \omega\left[u(C, z) - \overline{U} \ge 0\right] = 0 \\
&\dot{\pi}_K = \rho\pi_K - \pi_K F_K \\
&\dot{\pi}_X = \rho\pi_X - \pi_X G_X \\
&\dot{\pi}_C = \rho\pi_C + \pi_K - \omega u_C(C, z) = \pi_C\left[\rho + \frac{1}{\alpha}\right] + \pi_K \\
&\int_0^\infty e^{-\rho t}\frac{\partial L}{\partial \overline{U}}dt = 0
\end{aligned}
$$

The transversality conditions are

$$
\begin{aligned}
&\lim_{t\to\infty} \pi_C(t)e^{-\rho t}C(t) = 0 \\
&\lim_{t\to\infty} \pi_K(t)e^{-\rho t}K(t) = 0 \\
&\lim_{t\to\infty} \pi_X(t)e^{-\rho t}X(t) = 0
\end{aligned}
$$

We also make use of the Hamilton–Jacobi–Bellman equation

$$H = \rho V$$

which implies that

$$\pi_K[F(K, E) - C] + \pi_X[G(X) - E] + \pi_C[z] = 0 \tag{15.19}$$

We now show that equation (15.19), together with the above optimality conditions, implies a saving rule that must hold in order to satisfy the modified difference principle. Let us define P_X and γ by

$$
\begin{aligned}
P_X &= \frac{\pi_X}{\pi_K} = F_E \\
\gamma &= \frac{(-\pi_C)}{\pi_K}
\end{aligned}
$$

Then

$$\frac{1}{F_E}\frac{d}{dt}[F_E] = \frac{\dot{\pi}_X}{\pi_X} - \frac{\dot{\pi}_K}{\pi_K} = \rho - G_X - \rho + F_K = F_K - G_X$$

$$\ln\gamma = \ln(-\pi_C) - \ln\pi_K$$

$$\frac{\dot{\gamma}}{\gamma} = \frac{\dot{\pi}_C}{\pi_C} - \frac{\dot{\pi}_K}{\pi_K} = \left[\rho + \frac{1}{\alpha}\right] - \frac{1}{\gamma} - \rho + F_K = F_K + \frac{1}{\alpha} - \frac{1}{\gamma}$$

Thus

$$\dot{\gamma} = -1 + \gamma\left[\frac{1}{\alpha} + F_K\right] \tag{15.20}$$

Integration yields

$$\gamma(t) = \int_t^\infty e^{-\int_t^\tau \frac{1}{\alpha} + F_K(s)ds} d\tau$$

Substituting γ into equation (15.19), we get

$$\dot{K} + F_E[G(X) - E]E = \gamma\dot{C} = \dot{C}\int_t^\infty e^{-\int_t^\tau \frac{1}{\alpha} + F_K(s)ds} d\tau$$

If the production function is Cobb-Douglas and is independent of X, and $G(X) = 0$, one obtains the following system of five differential equations:

$$\dot{K} - F_E E = \gamma\dot{C}$$

$$\dot{\gamma} - \gamma\left(\frac{1}{\alpha} + F_K\right) = -1$$

$$\beta\frac{\dot{K}}{K} + (\sigma - 1)\frac{\dot{E}}{E} = \beta K^{\beta-1}E^\sigma$$

$$\dot{X} = -E$$

$$\alpha\dot{C} + C = \overline{Q} \equiv U^{-1}(\overline{U})$$

This system of five differential equations (in the five variables K, E, C, X, γ, and the unknown $\overline{U}$) can be solved using the following six boundary conditions:

$$X(0) = X_0$$
$$X(\infty) = 0$$
$$K(0) = K_0$$
$$\lim_{t\to\infty} K(t) = \infty$$
$$\lim_{t\to\infty} E(t) = 0$$
$$\lim_{t\to\infty} C(t) = U^{-1}(\overline{U})$$

In principle, one can then show that the maximin welfare level is an increasing function of the initial stocks K_0 and X_0:

$$\overline{U} = \overline{U}(K_0, X_0)$$

Starting with $X_0 > 0$, along the maximin welfare path, consumption is rising ($\dot{C} > 0$). But $\dot{C}$ tends to zero as t tends to infinity.

Appendix II: Proof of Proposition 5

Case A: $K_0 < \widehat{K}$

As is well known, if $\lambda = 0$ then the optimal path is the lower stable branch of the saddlepoint $(\widehat{K}, \widehat{C})$ where $\widehat{C} = F(\widehat{K}) - \delta\widehat{K}$. Let this branch be represented by the closed-loop control rule $C(t) = \phi^u(K(t))$ for $K < \widehat{K}$ where the superscript u denotes the fact that the solution corresponds to the standard utilitarian solution. Clearly, $\phi^u(K) > F(K) - \delta K$ for all $K < \widehat{K}$. If $\lambda > 0$ then the rule $\phi^u(K)$ is no longer optimal, because if it were, then we would have $C(0) = \phi^u(K_0)$ and $C(t) > \phi^u(K_0)$ for all $t > 0$. Thus we would have $\mu(t) = 0$ for all $t > 0$. This would violate (15.15).

If $\lambda = 1$, the optimal solution is $C(t) = F(K_0) - \delta K_0$ for all $t \geq 0$.

Consider the following conjectured optimal path for $\lambda \in (0, 1)$. The path consists of two phases. In Phase I, over some initial time interval $[0, t_\lambda]$, $C(t) = C_\lambda$ (a constant) where

$$F(K_0) - \delta K_0 > C_\lambda > \phi^u(K_0)$$

During this phase, $\dot{K}$ grows at the rate

$$\dot{K} = F(K) - \delta K - C_\lambda$$

until the time t_1 is reached, at which $K(t_1) = K_\lambda$ where

$$\phi^u(K_\lambda) = C_\lambda$$

During this phase,

$$[1 - \lambda + \mu]\, U'(C_\lambda) = \psi$$

Thus

$$\frac{\dot{\mu}}{1 - \lambda + \mu} = \frac{\dot{\psi}}{\psi} = \rho + \delta - F'(K)$$

In Phase II, choose $C(t) = \phi^u(K(t))$. During this phase, the standard utilitarian solution prevails.

To find C_λ and t_λ, one proceeds as follows. For any arbitrary pair (C_λ, t_λ), consider the solution of the differential equations $\dot{K} = F(K) - \delta K - C_\lambda$ and $\dot{\mu} = [\rho + \delta - F'(K)](1 - \lambda + \mu)$ with the boundary conditions $K(0) = K_0$ and $\mu(t_\lambda) = 0$. The solution yields μ_0 and $K(t_\lambda)$. These two numbers (which depends on the arbitrary pair (C_λ, t_λ)) must satisfy the following requirements

$$\int_0^{t_\lambda} [\rho\lambda - \mu]\, e^{-\rho t} dt = 0$$

$$\phi^u(K(t_\lambda)) = C_\lambda$$

These two conditions determine (C_λ, t_λ). The path thus constructed satisfies all the necessary and sufficient conditions.

Case B: $K_0 > \widehat{K}$

In this case, if $\lambda = 0$, the solution would be to follow the upper stable-branch of the utilitarian saddlepoint $(\widehat{K}, \widehat{C})$. Let this branch be represented by the closed-loop control rule $C(t) = \sigma^u(K(t))$ for $K > \widehat{K}$ where the superscript u denotes the fact that the solution corresponds to the standard utilitarian solution. Clearly, $\sigma^u(K) < F(K) - \delta K$ for all $K > \widehat{K}$. If $\lambda > 0$ then the rule $\sigma^u(K)$ is no longer optimal because we would have $\mu(t) = 0$ for all t, which violates (15.15).

Consider the following conjectured optimal path for $\lambda \in (0, 1)$. The path consists of two phases. The second phase, which begins at some time $T > 0$, consists of stationary consumption at some level $\widetilde{C}_\lambda$ where

$$\widehat{C} < \widetilde{C}_\lambda < C_G$$

where C_G is the golden-rule consumption level corresponding to the golden rule capital stock level K_G, i.e. $F'(K_G) = \delta$. Let $\widetilde{K}_\lambda$ be defined by

$$F(\widetilde{K}_\lambda) - \delta \widetilde{K}_\lambda = \widetilde{C}_\lambda$$

Then

$$\rho + \delta > F'(\widetilde{K}_\lambda) > \delta$$

During this phase, capital is also stationary at $K = \widetilde{K}_\lambda$

$$\frac{\dot{\mu}}{1 - \lambda + \mu} = \frac{\dot{\psi}}{\psi} = \rho + \delta - F'(\widetilde{K}_\lambda) > 0$$

Using (15.16), we have

$$\lambda = \int_T^\infty \mu e^{-\rho t} dt$$

Thus, for a given $\widetilde{K}_\lambda$, and noting that $\mu(T) = 0$, this integral determines the time T as a decreasing function of $\widetilde{K}_\lambda$ and hence of $\widetilde{C}_\lambda$

$$T = T(\widetilde{C}_\lambda)$$

Consider now the phase I, which is utilitarian. For given K_0, there exists a unique trajectory in the (K, C) phase diagram that leads to $(\widetilde{K}_\lambda, \widetilde{C}_\lambda)$. This trajectory takes $T^\#$ units of time.

$$T^\# = T^\#(\widetilde{C}_\lambda)$$

If $T^\#(\widetilde{C}_\lambda)$ is an increasing function of $\widetilde{C}_\lambda$, then there is at most one $\widetilde{C}_\lambda$ at which $T = T^\#$.

Notes

1. In his essay, 'Idea for a Universal History with a Cosmopolitan Purpose,' Kant put forward the view that nature is concerned with seeing that man should work his way onwards to make himself worthy of life and well-being. He added: 'What remains disconcerting about all this is firstly, that the earlier generations seem to perform their laborious tasks only for the sake of the later ones, so as to prepare for them a further stage from which they can raise still higher the structure intended by nature; and secondly, that only the later generations will in fact have the good fortune to inhabit the building on which a whole series of their forefathers... had worked without being able to share in the happiness they were preparing.' See Reiss (1970, p. 44).
2. Rawls (1999, p. 253).
3. The earliest model of optimal savings was formulated by Ramsey (1928), who thought it would be is unethical to discount the utility of future generations. There was a burst of activities on optimal savings theory, beginning in the 1960s. See Sen (1961), Tobin (1966, Chapter 9), Koopmans (1965), Solow (1970, Ch. 5), Arrow and Kurz (1970), among others. Amongst later contributions with emphasis on welfare significance in a dynamic economy, a (biased)

sample would include Weitzman (1976), Kemp and Long (1982, 1998), Beltratti, Chichilnisky and Heal (1994, 1995), Chichilnisky (1996), Asheim (2000). For some non-utilitarian contributions and papers with emphasis on constant consumption, see Solow (1974, 1986), Dixit, Hammond and Hoel (1980), Mitra (1983), Hartwick and Long (1999), Asheim, Buchholtz and Withagen (2003).

4. David Hume, Adam Smith, and John Stuart Mill were among the greatest utilitarian social theorists cum economists.
5. Rawls (1999, p. 10).
6. Rawls (1999, p. 53).
7. Rawls (1999, p. 72).
8. Rawls (1999, pp. 67–8). The tension between equity and efficiency noted by Rawls was well recognized by practitioners. As Lee Kwan Yew stated, 'If performance and rewards are determined by the marketplace, there will be a few big winners, many medium winners, and a considerable number of losers. That would make for social tensions because a society's sense of fairness is offended.... To even out the extreme results of free-market competition, we had to redistribute the national income...If we over-redistributed by higher taxation, the high performers would cease to strive. Our difficulty was to strike the right balance' (Lee, 2000, p. 116).
9. A more elaborate form of the difference principle is called the 'lexical difference principle' , see Rawls (1999, p. 72), and Sen (1976).
10. Solow acknowledged that in applying the maximin rule to intergenerational equity, he is 'plus rawlsien que Rawls'.
11. Rawls (1999, p. 254).
12. Rawls (1999, p. 10).
13. Rawls (1999, p. 9).
14. Strictly speaking, the utilitarian principle does not require interpersonal comparison of levels of utility: only increments of individual utility are compared. In contrast, Rawls' difference principle does not require interpersonal comparison of increments, but requires interpersonal comparison of utility levels. See Dasgupta (1974a), Sen (1976).
15. It has been argued that, contrary to Rawls' opinion, the parties behind the veil of ignorance would choose the utilitarian principle. The arguments advanced by Harsanyi (1953) support the utilitarian principle, though in a slightly different context. In contrast, Strasnik (1976) showed that Rawls's difference principle can be derived from axioms which are "reasonable" under some interpretation of the original position. See also Epstein (1986a,b), Chichilnisky (1996), Suga and Udagawa (2004, 2005) for some axiomatic approaches to intergenerational equity.
16. Does a reflective equilibrium exist? If so, is it unique? (And is it locally stable?) Rawls (1999, p. 44) believes 'it would be useless to speculate about these matters here.'
17. A rational person has a 'coherent set of preferences between the options open to him. He ranks these options according to how well they further his purposes' (Rawls, 1999, p. 124). It is also assumed that he does not suffer from envy. (Rawls, 1999, p. 124).
18. Rawls (1999, p. 118).
19. Rawls acknowledges that 'it may turn out, once the veil of ignorance is removed, that some of them for religious or other reasons may not, in fact, want more of these goods' (Rawls, 1999,p. 123).

20. Rawls (1999, p. 122).
21. Rawls (1999, p. 111).
22. Rawls (1999, p. 255).
23. Rawls (1999, p. 254).
24. Rawls (1999, p. 255).
25. Rawls (1999, p. 37).
26. Rawls (1999, p. 255).
27. For an earlier version of this model, see Long and Yang (1999).
28. Rawls (1999, p. 121).
29. Rawls (1999, p. 121).
30. Rawls (1999, p. 121).
31. Rawls (1999, p. 256).
32. Rawls (1999, p. 256).
33. Rawls (1999, p. 257).
34. Rawls (1999, p. 255).
35. Rawls (1999, pp. 257–8).
36. Wilde (1926, Act 3, p. 113).
37. Rawls (1999, p. 256).
38. Arrow (1973, Theorem 2) showed that the maximin programme can display a saw-tooth shaped path of consumption and utility. This was generalized by Suga (2004).
39. See Dasgupta (1974b), Kohlberg (1976), Lane and Mitra (1981), Leininger (1986), Bernheim and Ray (1983, 1987), Pezzey (1994).
40. For an axiomatization of utilitarianism and egalitarianism, see Epstein (1986a,b). For an axiomatization of the maximin objective, see Strasnik (1976).

References

Arrow, K. J. and M. Kurz (1970) *Public Investment, the Rate of Return and Optimal Fiscal Policy* (Baltimore: Johns Hopkins University Press).

Arrow, K. J. (1973) 'Rawls' Principle of Just Saving', *Swedish Journal of Economics*, vol. 75, pp. 323–35.

Asheim, Geir B. (1988), 'Rawlsian Intergenerational Justice as a Markov-Perfect Equilibrium in a Resource Technology', *Review of Economic Studies*, vol. 55(3), pp. 469–83.

Asheim, G. B. (2000) 'Green National Accounting: Why and How?', *Environment and Development Economics*, vol. 5(1), pp. 25–48.

Asheim, G. B., W. Buchholtz and C. Withagen (2003) 'The Hartwick Rule: Myths and Facts', *Environmental and Resource Economics*, vol. 25(2), pp. 129–50.

Beltratti, A., G. Chichilnisky and G. Heal (1994) 'The Environment and the Long Run: A Comparison of Different Criteria', *Ricerche Economiche*, vol. 48, pp. 319–40.

Beltratti, A., G. Chichilnisky and G. Heal (1995) 'Sustainable Growth and the Green Golden Rule', in I. Goldin and A. Winters (eds) *The Economics of Sustainable Development* (Cambridge: Cambridge University Press).

Bernheim, B. D. and D. Ray (1983) 'Altruistic Growth Economies: I. Existence of Bequest Equilibria', Technical Report Number 419, Institute for Mathematical Studies in the Social Sciences, Stanford University.

Bernheim, B. D. and D. Ray (1987), ' Growth with Altruism', *Review of Economic Studies*, vol. 54, pp. 227–43.

Burmeister, E. and P. Hammond (1978) 'Maximin Paths of Heterogeneous Capital Accumulation and the Instability of Paradoxical Steady States', *Econometrica*, vol. 45(4), pp. 853–70.

Calvo, G. (1978) 'Some Notes on Time Inconsistency and Rawls' Maximin Criterion', *Review of Economic Studies*, vol. 45, pp. 97–102.

Chichilnisky, G. (1996) 'An Axiomatic Approach to Sustainable Development', *Social Choice and Welfare*, vol. 13 (3), 231–57.

Dasgupta, P. (1974a) 'Some Problems Arising from Professor Rawls' Conception of Distributive Justice', *Theory and Decision*, vol. 4, pp. 325–44.

Dasgupta, P. (1974b) 'On Some Alternative Criteria for Justice Between Generations', *Journal of Public Economics*, vol. 3, pp. 405–23.

Dixit, A., P. Hammond and M. Hoel (1980) 'On Hartwick's Rule for Regular Maximin Paths of Capital Accumulation and Resource Depletion', *Review of Economic Studies*, vol. 47, pp. 551–6.

Dockner, E., S. Jorgensen, N. V. Long, and G. Sorger (2000) *Differential Games in Economics and Management Science* (Cambridge: Cambridge University Press).

Epstein, L. G. (1986a) 'Intergenerational Preference Orderings', *Social Choice and Welfare*, vol. 3, pp. 151–60.

Epstein, L. G. (1986b) 'Intergenerational Consumption Rules: An Axiomatization of Utilitarianism and Egalitarianism,' *Journal of Economic Theory*, vol. 38, pp. 280–97.

Harsanyi, J. C. (1953) 'Cardinal Utility in Welfare Economics and in the Theory of Risk-taking', *Journal of Political Economy*, vol. 61(5), pp. 434–7.

Hartwick, J. M. and N.V. Long (1999) 'Constant Consumption and Economic Depreciation of Natural Capital', *International Economic Review*, vol. 40(1), pp. 53–62.

Kemp, M. C., and N. V. Long (1982) 'On the Evaluation of National Income in a Dynamic Economy', in G. Feiwell (ed.), *Samuelson and Neoclassical Economics* (Boston: Marinus Nijhoff Publishing), pp. 185–98.

Kemp, M. C., and N. V. Long (1998) 'On the Evaluation of National Income in a Dynamic Economy: Generalization', in K. Jaeger and K.J. Koch (eds), *Trade, Growth, and Economic Policies in Open Economies* (Berlin: Springer Verlag).

Kohlberg, E. (1976) 'A Model of Economic Growth with Altruism Between Generations', *Journal of Economic Theory*, vol. 13, pp. 1–13.

Koopmans, T. C. (1965) 'On the Concept of Optimal Economic Growth', *Pontificiae Academiae Scientiarum Scripta Varia*, vol. 28, pp. 225–87.

Lane, J. and T. Mitra (1981) 'On Nash Equilibrium Programmes of Capital Accumulation under Altruistic Preferences', *International Economic Review*, vol. 22, pp. 309–31.

Lee, K. Y. (2000) *From Third World to First: The Singapore Story 1965–2000. Memoirs of Lee Kwan Yew* (Singapore: Singapore Press Holdings).

Leininger, W. (1985) 'Rawls' Maximin Criteria and Time-Consistency: Further Results', *Review of Economic Studies*, vol. 52, pp. 505–13.

Leininger, W. (1986) 'The Existence of Perfect Equilibria in a Model of Growth with Altruism between Generations', *Review of Economic Studies*, vol. 53, pp. 349–67.

Leonard, D. and N. V. Long (1992) *Optimal Control Theory and Static Optimization in Economics* (Cambridge: Cambridge University Press).

Long, N. V. and Z. Yang (1999) 'Constant Utility with Growing Consumption: A New Rawlsian Approach', Ch. 2, PhD Thesis, McGill University, 'Three Essays on Sustainable Growth and Environmental Control' (Library Ref. AS42 M3 1999 37).

Mitra, T. (1983) 'Limits on Population Growth under Exhaustible Resource Constraint', *International Economic Review*, vol. 24, February, pp. 155–68.

Pezzey, J. (1994) 'Theoretical Essays on Sustainability and Environmental Policy', PhD Thesis, University of Bristol.

Ramsey, F. P. (1928) 'A Mathematical Theory of Saving', *Economic Journal*, vol. 38, pp. 543–59.

Rawls, J. (1958) 'Justice as Fairness', *The Philosophical Review*, vol. 67(3), pp. 164–94.

Rawls, J. (1963) 'The Sense of Justice', *The Philosophical Review*, vol. 72(2), pp. 281–305.

Rawls, J. (1971) *A Theory of Justice* (Cambridge, MA: The Belknap Press of the Harvard University Press).

Rawls, J. (1999) *A Theory of Justice* (revised edition) (Cambridge, MA: The Belknap Press of the Harvard University Press)

Reiss, H. (1970) *Kant's Political Writings* (Cambridge: Cambridge University Press)

Rodriguez, A. (1981) 'Rawls' Maximum Criterion and Time Consistency: A Generalization', *Review of Economic Studies*, vol. 48, pp. 599–605.

Sen, A. K. (1961) 'On Optimizing the Rate of Saving', *Economic Journal*, vol. 71, pp. 479–96.

Sen, A. K. (1976) 'Rawls versus Bentham: An Axiomatic Examination of the Pure Distribution Problem', in N. Daniels (ed.), *Reading Rawls: Critical Studies on Rawls' A Theory of Justice* (New York: Basic Books).

Solow, R. M. (1970) *Growth Theory* (New York: Oxford University Press).

Solow, R. M. (1974) 'Intergenerational Equity and Exhaustible Resources', *Review of Economic Studies* (Symposium issue), vol. 41, pp. 29–45.

Solow, R. M. (1986) 'On the Intergenerational Allocation of Natural Resources', *Scandinavian Journal of Economics,* vol. 88, pp. 141–9.

Strasnik, S. (1976), 'Social Choice and the Derivation of Rawls's Difference Principle', *The Journal of Philosophy*, vol. 73(4), pp. 85–99.

Suga, Koichi (2004) 'On the Role of Externalities in the Arrow–Dasgupta Economy', *Waseda Economic Papers*, no. 43, pp. 1–24.

Suga, K. and D. Udagawa (2004), 'Axiomatic Characterization of the Maximin Principle in the Arrow–Dasgupta Economy', School of Political Science and Economics, Waseda University.

Suga, K. and D. Udagawa (2005) 'Characterization of the Maximin Choice Function in a Simple Dynamic Economy', School of Political Science and Economics, Waseda University.

Tobin, J. (1966) *National Economic Policy* (New Haven: Yale University Press).

Weitzman, M. (1976) 'On the Welfare Significance of National Product in a Dynamic Economy', *Quarterly Journal of Economics*, vol. 90(1), pp. 55–68.

Wilde, O. (1926) *Lady Windermere's Fan* (London: Methuen and Co.).

Withagen, C. and G. B. Asheim (1998) 'Characterizing Sustainability: the Converse of Hartwick's Rule', *Journal of Economic Dynamics and Control,* vol. 323(1), pp. 159–65.

16

Normative Approaches to the Issues of Global Warming: Responsibility and Compensation*

Kotaro Suzumura and Koichi Tadenuma

16.1 Introduction

The study of intergenerational equity may be broken down into four broad categories in accordance with the following two simple questions. The first is whether we are merely concerned with the equity issues among generations that are adjacent or overlapping along the time axis, or whether our concern about intergenerational equity is broad enough to cover even distant future generations (which we would never encounter in our own lifetimes). The second is related to informational bases of intergenerational equity. It asks what information is relevant to judgements of intergenerational equity, or whether we should assume that only utility or welfare information is relevant to equity judgements. Among the four categories identified by the combination of answers to these two simple questions, this chapter focuses on the analysis of equity among distant generations, without assuming that only utility or welfare information is relevant.

To crystallize the focal issues of intergenerational equity that we analyze in this chapter, consider the basic nature of the problem of global warming. The accumulation of greenhouse gases generated from economic activity may exert a persistent influence on the global climate on a scale that we have never

* We are most grateful to Marc Fleurbaey, Ko Hasegawa, Hajime Hori, Susumu Morimura, Koichi Suga and Makoto Usami for their helpful comments at the Hakone Conference on Intergenerational Equity. Our special gratitude is due to Peter Vallentyne, who gave us many detailed comments and suggestions which greatly helped us improve the content as well as exposition of this chapter. Financial support was provided by the Ministry of Education, Culture, Sports, Science and Technology of Japan through the Grant-in-Aid for Scientific Research in Priority Areas and the 21st Century Center of Excellence Program on the Normative Evaluation and Social Choice of Contemporary Economic Systems. Needless to say, we are solely responsible for any remaining deficiency of the final product.

encountered in the past. As the accumulated greenhouse gases may stay in the air a long time, the current emission of greenhouse gases may exert persistent effects on many generations, even in the distant future. The complexity of this problem cannot be exaggerated. Observe that everybody leading a normal economic life is unable to avoid generating greenhouse gases; thus, everybody cannot but be partially responsible for causing the problem of global warming. Furthermore, the full extent of the concerned parties goes far beyond the present generation. Not only the present economic activities but also the past economic activities at least since the Industrial Revolution have contributed to the accumulation of greenhouse gases. But most of the past generations which should be held responsible for the past accumulation of greenhouse gases no longer exist. If we look ahead along the historical time axis into the indefinite future, those generations that would be most affected by global warming would most likely be generations born in the distant future. However, these future generations simply do not as yet exist, and we are left with no information about their population size and human character. We are thus confronted with the serious problem of environmental externalities, a large part of whose culprits no longer exist, and a large part of whose victims do not yet exist. In between these non-existing past generations and non-existing future generations lies the present generation that is only partially responsible for the problem at hand.

If the problem of global warming should be kept under control for the sake of equitable treatment of distant future generations, it is clear that the present generation is in a unique position to take action. But the reason why this action, which may require serious sacrifices and strenuous efforts on the part of those who take action, should be taken single-handedly by the present generation surely requires some rational explanation. This chapter represents a modest attempt to seek such a rational explanation. In what follows, our attempt will be based on a simple principle which is relatively new – at least in its contemporary resurgence. It is going to play a crucial role in answering the following two imaginary questions that are supposed to be raised by the present generation:

(a) Should we care about the possible plight of future generations? Those who might be most seriously affected by global warming would emerge after we are all gone, and we might never be forced to confront their plight.
(b) Should we be held single-handedly responsible? There are many past generations that should at least in part share the blame for what we are jointly responsible. True enough, scientific evidence as well as public awareness of the problem of global warming is much more solid now than it used to be. But should we be held single-handedly responsible for fixing the problems created by past generations?

It is for the purpose of coping with these problems that we invoke the principle of responsibility and compensation. The origin of this principle may be traced back all the way to Aristotle, and its modern resurgence is due to the political philosopher, Ronald Dworkin (1981a, 1981b, 2000). In essence, it asserts that one should be held responsible for the culmination outcome of one's own voluntary choice, over which one has a full power of control, but not necessarily otherwise. However, given the nature of the problem we face in the context of global warming, we need some further reasoning and scaffolding in order to make this principle applicable to the problem at hand. It is hoped that this modified principle of responsibility and compensation would help us understand the unilateral duty of the present generation in the face of global warming. At the very least, it is our hope that our modest attempt in this chapter would motivate us to think rationally about the very serious ethical problem we are jointly facing.

The structure of the rest of this chapter is as follows. In order to set the stage for our reasoning, section 16.2 is devoted to clarifying the structure of global environmental problems such as global warming, and section 16.3 discusses the fundamental non-identity problem posed by Derek Parfit (1984) in the specific context of global warming. Section 16.4 then examines the effectiveness of some normative criteria that have been invoked by traditional welfare economics in the arena of environmental externalities. After these preliminary steps, section 16.5 proposes a fundamental normative principle, to be called the *principle of responsibility for selecting a future path*, which is a modified version of Dworkin's principle of responsibility and compensation in the present context, and argues that the primary responsibility of the present generation is to choose a future path which is accountable as socially rational in accordance with some coherent principle. Section 16.6 adopts the principle of responsibility and compensation as our evaluation criterion of future paths with the purpose of extracting the implications thereof for the choice of a future path. Section 16.7 discusses the implication of the same principle in the context of intergenerational burden sharing of the cost of abating global warming. Section 16.8 concludes this chapter by making several final observations.

16.2 Temporal structure of the problems of global warming

Global warming is an example of environmental externalities – that is, where the actions of some economic agents unintentionally and incidentally affect the payoffs of some other economic agents without being mediated by market mechanisms. However, there are several conspicuous features of the problem that make it unique among many problems of environmental externalities. At the risk of slight overlap with what we have already mentioned, some of these features will be reiterated in this section with the purpose of bringing its uniqueness into clear relief.

Non-coexistence of the culprits and the victims

In the case of environmental disruption such as water contamination by emission from factories, or noise in the neighbourhood of an airport, the culprits and the victims of detrimental externalities usually coexist. In sharp contrast, not all agents involved in the problem of global warming coexist. Indeed, those generations that would be most severely affected by greenhouse gas emissions come into existence only much later. Thus, the majority of culprits of the past emissions and the major victims from the resulting global warming do not exist at present. Besides, those who can possibly participate in any attempt to cope with the problem of global warming consist solely of the present generation that, however, accounts for only a tiny fraction of the culprits as a whole. Furthermore, it is difficult, to say the least, to represent the legitimate claims of distant future generations in the present social decision-making procedure when the population size and human identity of these potential people are not known.

Non-limitation of the culprits and the victims

In most cases of environmental disruptions, those who trigger the problem and those who suffer from the problem coexist and are limited in number. These features are prerequisites for the traditional resolution schemes to be applicable to these cases of environmental disruptions. In the case of global warming, however, neither the culprits nor the victims are limited in time and/or space.

First, *the culprits of global warming are not limited in space*. The emission of greenhouse gases is unavoidable in every normal economic activity, so that every human being cannot but be partially responsible for global warming. It follows that, to control the emission of greenhouse gases effectively and equitably, cooperation among people in all countries and regions is necessarily called for.

Secondly, *the culprits of global warming are not limited in time*. The problem of climate change at each time may serve as a link between many past generations and many future generations, where the intertemporal linkage may be extended into the indefinite future, and the interregional linkage may be extended to almost everywhere on the earth. Thus, it is not only the current economic activities, but also the past economic activities ever since the Industrial Revolution, that should be counted in the factors that triggered global climate change.

Thirdly, *the victims of global warming are not limited in space*. Climate change at any historical time cannot but affect the living standard of people, no matter where they live on the earth. Finally, *the victims of global warming are not limited in time*. The emission of greenhouse gases at any historical time may exert influence on the standard of living of indefinite future generations.

We see later that these features of global warming make it difficult for us to apply standard resolution schemes for environmental externalities proposed in the literature of economics.

16.3 Non-identity problem in the context of global warming

To summarize our argument so far, global warming is an externality problem over extremely long periods; people involved in the problem will range over many generations; they do not coexist, but only successively appear and/or disappear along the historical time path. This is the basic structure of the problem to be kept in mind, but there is another conspicuous feature to be emphasized.

Generally speaking, among infinitely many potential historical paths of human evolution, the one that describes our actual historical path from the past to the present is already fixed, as are the people who have existed and presently exist along the realized historical path. However, what type of people will emerge in the future, and in what numbers will they emerge, depend on the actions taken by the present generation, and are indeterminate at the point of decision-making. Let us provide some concrete examples to show why this is the case.

(1) Let us compare the following two alternative scenarios: (a) to adopt policy measures that strictly limit the use of petroleum in developed countries; and (b) to impose no restrictions whatsoever. Needless to say, the use of petroleum plays such a crucial role in all aspects of human life that the styles of food, clothes and shelter, and convenience and opportunity to travel would all differ immensely depending on which of the two alternative scenarios was chosen and implemented. Depending on the choice made between these two options, people would meet and marry different partners, have different families and lifestyles, and accumulate different life experiences. Thus, the culmination outcomes over several generations would most probably be that a different set of people would exist.

(2) The total amount of greenhouse gas emissions would depend not only on the size of the population, but also on the per head greenhouse gas emissions by the given population. Thus, there would probably be very different culmination outcomes with respect to the scale as well as characteristics of future generations depending on whether we would choose and implement policy measures to control population explosions in developing countries or not.

(3) If greenhouse gas emissions were not controlled, some island countries in the Pacific might be submerged under water, and the regional distribution of population would be changed immensely. Likewise, population and its regional distribution could be affected to a large extent by the possibility

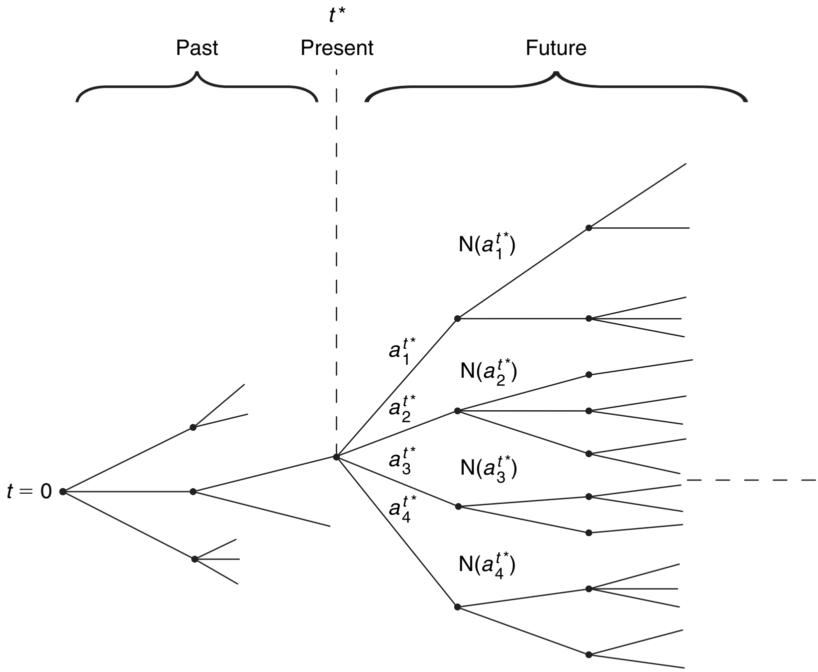

Figure 16.1 Past and future time paths

that the areas so far categorized to be permanently frozen become suitable for cultivation due to global warming.

The important point is that which people exist is in fact malleable in response to the choice of actions by the present generation. Figure 16.1 illustrates this fact graphically. The path from the past to the present is uniquely determined. By contrast, numerous paths from present onward are possible, which are contingent on the actions chosen by the present generation as well as the future generations. Let the 'present' be time t^*. Then, whether a certain action is open for choice at time t^* depends on the path of actions $\mathbf{a}^{t^*-1} = (a^0, \ldots, a^{t^*-1})$ realized from the starting point of history at 0 until t^*-1. The set of all possible actions at time t^* is denoted $A^{t^*}(\mathbf{a}^{t^*-1})$. When the present generation chooses an action $a^{t^*} \in A^{t^*}(\mathbf{a}^{t^*-1})$, the set of people who may possibly exist after t^* is labelled $N(a^{t^*})$. Those belonging to the set $N(a^{t^*})$ are all people who exist on the paths following the branch decided by the action a^{t^*}. They may be called 'potential people relative to time t^*'. In general, if it is the case that $a^{t^*} \neq b^{t^*}$, then $N(a^{t^*}) \neq N(b^{t^*})$. This is the *non-identity problem* for the future generations pointed out by Derek Parfit (1981, 1984).

Observe that Parfit's claim is that, depending on the path to be followed, the individuals are not the same. This is not confined to the difference in their biological features. Human beings are social as well as biological entities. Thus, not only biological attributes, but also various social attributes such as preferences, abilities to consume and work, abilities to understand, communicate, collaborate with other people, and abilities to make sensible judgements should all be considered as important factors, in terms of which individuals are to be judged whether or not they remain to be the same individuals. It is clear that preferences and abilities are characteristics that are formed endogenously through social interactions over a long period. Furthermore, the effectiveness of a specific ability differs substantially from one social environment to the other. If one considers that the counter-measures for global warming would change social structures and economic environments over a long period, one cannot but find it unavoidable that the identity of future people would be substantially different in accordance with the choice of various measures against global warming.

Let us go back, then, to the global warming problem, where the generations have dual aspects, that is, a passive aspect and an active aspect. On the one hand, they are affected by climate changes caused by the accumulation of greenhouse gases emitted before their birth. On the other hand, they themselves make decisions on the control of greenhouse gas emission, which in turn exert influences on the living standard of the generations born after them. Depending upon which aspect of generations we choose for scrutiny, the non-identity problem has two distinct implications. In the first place, the preferences of the future generations depend crucially upon the policies to be implemented. In the second place, the standard of value judgements of the future generations, in terms of which they themselves would decide on the policies to control greenhouse gases at their own decision time, are also contingent on the choices made by the present generation.

As we have already pointed out, in order to keep climate change under control at any future point of time, it is necessary to control greenhouse gas emissions by many preceding generations (*non-limitation of the culprits in time*). Although each future generation makes its own decision autonomously, their standard of value judgements cannot but be formed at least in part under the influence of the present generation. To this extent, the policy choice of the future generation cannot but depend upon the behaviour of the present generation. Thus, it is by no means straightforward to answer which policy at present is more desirable for the future generations than others. Indeed, we must take into account not only the direct effect of each policy on the global climate of the earth (the *direct effect*), but also the indirect effect on the preferences of future generations (the *preference effect*), as well as on the standard of value judgements of the future generations (the *value standard effect*). This complexity is a logical consequence of a non-identity problem. The Pandora's Box opened by Parfit poses a serious problem in the analysis of global warming.

16.4 Effectiveness of orthodox economic analyses

Confronted with the problem of global warming, how should we determine the distribution of living standards among successive generations as well as different groups of people within the same generation? Standard welfare economics developed various criteria for judging social desirability of one resource allocation vis-à-vis the other. In this section, we examine the effectiveness of these normative criteria in the context of intergenerational resource allocations in the presence of global warming.

Pareto criterion and the compensation principles

If we strictly follow the 'new' welfare economics that excludes interpersonal comparisons of well-being, the only normative standard we may use in comparing alternative economic policies is the Pareto criterion. Suppose that we must choose between the culmination outcome of one policy **a** and that of another policy **b** on the basis of preferences of all the persons involved. If no one prefers the outcome of **b** to the outcome of **a**, and at least some person prefers the outcome of **a**, the policy **a** is said to be Pareto superior to the policy **b**.

Since most policies involve conflict of interests among concerned persons, it is quite rare that one policy can be judged Pareto superior to another. However, if there is a common numeraire called 'money,' and if it is possible to transfer money from the persons who benefit to those who lose, then the applicable range of the Pareto criterion can be expanded through the *hypothetical* payment of compensations. Nicholas Kaldor, John Hicks, Tibor Scitovsky and Paul Samuelson developed their respective versions of a hypothetical compensation principle, upon which the cost–benefit analysis in applied welfare economics is based.

Are the Pareto criterion and the compensation principles useful in judging on alternative policies against global warming? Because any policy to abate the emission of greenhouse gases necessarily reduces the level of economic activities, it is certain that the present generation would have to incur some losses. Therefore, in order for the Pareto criterion and the compensation principles to be applicable to the evaluation of the policy against global warming, a transfer of money must be conceivable from the distant future generations, which would be the beneficiaries of the policy, to the present generation. However, such a monetary transfer is hard to visualize due to the two features of the problem of global warming: non-coexistence of the culprits and the victims and non-limitation of the culprits and the victims in space. Because the culprits (the present generation) and the victims (the distant future generations) do not exist simultaneously, the monetary transfers must be carried out over distant time. If all the concerned persons lived in a specific country, then the transfers of money would be easy: the persons at present issue bonds outside the country, and have the persons in the distant future take responsibility for redeeming the foreign bonds. In this way, intertemporal monetary transfers from the distant future to the present would be made possible. However, there is no national border in the problem of global warming. Because all those who belong to the present generation would get some losses from implementing policies to abate global warming, and all those who belong to the distant future generations would benefit from it, it is simply impossible to use foreign bonds as a device for transferring compensatory payments from the distant future generations to the present generation. The only method to carry out compensatory payments would be to decrease the level of social capital that the present generation should leave for the distant future generations, and to increase the present consumption in exchange. However, it is most likely that the shift of resources from social capital to present consumption increases the present emissions of greenhouse gases. If compensatory monetary transfers are not feasible, then the Pareto criterion and the compensation principles lose their applicability to the task of judging whether a policy for abatement of global warming should be carried out or not.

Let us now call in Parfit's non-identity problem. Then, in the context of the problem of global warming, the Pareto criterion and the compensation principles are almost totally useless for choosing among alternative policies. Indeed, when we compare the policies a^{t^*} and b^{t^*}, the potential people who will exist in the future, $N(a^{t^*})$ and $N(b^{t^*})$, are such that $a^{t^*} \neq b^{t^*}$ implies $N(a^{t^*}) \neq N(b^{t^*})$. Since $N(a^{t^*})$ and $N(b^{t^*})$ consist of different people, who may not even coexist on the earth, neither the Pareto principle nor the various versions of the compensation principles make any coherent sense.[1]

Solution through negotiations by the concerned agents

It was Ronald Coase (1960) who proposed a well-known resolution to the problem of externalities, which is based on the direct negotiations between the culprits and the victims of negative environmental externalities. His proposition, known as the 'Coase theorem', states that, when the negotiation costs can be ignored, the externality problems are solved efficiently through negotiations of the concerned agents as long as it is clearly stipulated who are endowed with the property rights to begin with. Take, for example, the case of airport noise. If 'the right to live a quiet life' (or 'the right to produce noise without outside interference') is endowed *ab initio*, an agreement can be formed on the levels of noise as well as monetary compensation through negotiations between the concerned agents. This would lead to an efficient resolution of the problem of environmental externalities, although there would be significant difference in income distribution depending on how the rights are distributed in the first place.

In the context of the problem of global warming, it should be clear that the resolution through the direct negotiations in the spirit of Coase is impossible. This is because the time structure of the problem does not allow the culprits (those who emit greenhouse gases) and the victims (those who suffer from the resulting global warming) to exist simultaneously. It is true that, in a purely normative argument, we might assume a counterfactual situation in which all the concerned parties meet to negotiate. Even in the hypothetical setting such as this, however, we must admit that it is impossible to nominate rationally the agents who can properly represent the distant future generations if we accept Parfit's non-identity problem. Would they be the agents of people who would exist if some measures were taken at present against global warming, or those if the measures were not taken? In the former case, the people whom the agents represent would not exist when the policy measures against global warming were not implemented. Therefore, they could not be qualified to ask for compensation for the losses that would occur when no policy measures were implemented. In the latter case, the people whom the agents represent would not exist when the policy measures against global warming were implemented. Hence, the agents would become disqualified to represent the future generations and negotiate with the present generation once the present generation decided to implement such policy measures against global warming. In either case, it is fundamentally impossible to use rational negotiations between the concerned agents as the basis of normative theory in the context of global warming.

Rights and duties

Social justice is often expressed in terms of rights and duties. The common argument about 'the rights and duties between generations' in the context of global warming would be something like this: 'the future generations have the right to claim that the present generation should abate global warming, whereas the present generation has the corresponding duty owed to the future generations.' However, there are logical difficulties in articulating the relation between the future generations and the present generation in terms of such right-duty relation if we accept the non-identity problem.

Let the policy a^{t^*} mean 'implementing policy measures against global warming at time t^*', and the policy b^{t^*} mean 'not implementing policy measures against global warming at all.' In each case, the set of people who exist in the future are different: $N(a^{t^*})$ and $N(b^{t^*})$ have few, if any, individuals in common. Let us now ask: to which group of people, $N(a^{t^*})$ or $N(b^{t^*})$, does the right to request the present generation to restrain the progression of global warming belong? Or: to which group of people, $N(a^{t^*})$ or $N(b^{t^*})$, does the present generation owe the duty?

First, suppose that the group $N(b^{t^*})$, who would live in the world of progressed warming, has the right. However, once the present generation fulfills the duty corresponding to the exercise of the right by the future generations

$N(b^{t^*})$, the people that would exist in the future is the group $N(a^{t^*})$, and no longer $N(b^{t^*})$. The people of the group $N(b^{t^*})$ would never live in the world with restrained warming, and the exercise of the right means denying their own existence on the earth. John Stuart Mill (1859) once insisted that one cannot justify the slave contract to sell his own freedom under the name of the freedom of choice. Similarly, when one sets the right to decide one's own circumstances, it is irrational to set the right including the right to deny one's own existence.

On the other hand, there is no rational ground that the group $N(a^{t^*})$, who would exist in the world only with restrained warming, should have the right to demand that the present generation restrain global warming. If global warming were not restrained, then it is the group $N(b^{t^*})$ who would exist, and the group $N(a^{t^*})$ would not exist any longer. Hence, a similar argument parallel to the one given above would hold true. One possible counter-argument to this line of reasoning would be that the people who belong to the group $N(a^{t^*})$ should rather be entitled to claim a strong right against the present generation to restrain global warming because if it were not abated, the people $N(a^{t^*})$ would be denied of their very existence. However, it is possible to secure this right for the people in $N(a^{t^*})$ only by depriving a similar right from the people in $N(b^{t^*})$. There seems to be no sound ground that justifies such a discriminatory conferment of rights. It seems to us that the non-identity problem makes it impossible to base the counter-measures against global warming upon the duty of the present generation to the distant future generations.

In closing this section, let us call attention to the fact that we have carefully separated the argument that hinges on the acceptance of Parfit's non-identity problem, and the argument that does not. Given the relatively unfamiliar nature of Parfit's non-identity problem among economists, it is our belief that this separation would be useful in confirming the solidity of our argument.

16.5 Responsibility and compensation in the choice of future path

We have shown that it is impossible to find the normative basis of policy evaluation in the context of global warming by means of the Pareto criterion, the compensation principles, or the Coase paradigm of bilateral negotiations. We have also argued that it is impossible to base the policy measures against global warming on the rights of distant future generations. Should we be resigned in agnosticism on this issue? No. We must develop alternative normative principles in order to be freed from this trap of agnosticism.

As one of the elements of such a new theory, we would like to focus on the principle of responsibility and compensation, which has been introduced into welfare economics through recent studies by Ronald Dworkin (1981a, 1981b, 2000), John Roemer (1985, 1986), and Marc Fleurbaey (1995, 1998). What Fleurbaey (1998) christened 'responsibility by control' is the principle to the effect that one should be held responsible for the culmination outcome of one's own voluntary choice, over which one has a full power of control. In other words, the voluntary exercise of 'freedom of choice' generates responsibility on the part of a person in charge of the act of choice.

According to this principle, there is no justifiable reason for a person who developed special tastes for expensive cars to ask for 'compensation' from society, just because the degree of satisfaction of his desire is very low unless the expensive car is made available to him. The reason is that the low level of his satisfaction is caused by his expensive tastes, nourished by the exercise of his own free will. In such a case, the person himself should take responsibility for his own voluntary choice and should not shift the burden of his expensive tastes to society.

On the other hand, if a person is suffering from misfortune by birth, or social contingencies that cannot be attributed to his own voluntary choice, he should not be charged with personal responsibility for his plight. In this case, social 'compensation' must be paid for the losses caused by the non-responsible factors. Take, for example, an unfortunate person who became disabled due to an accident caused by someone else's drunken driving. It is clear that he should receive social compensation, because he is suffering from a loss caused by an uncontrollable factor, for which he should not be held personally responsible. It is true that the responsibility of compensatory payment to him should mostly belong to the individual who drove while drunk, but it is the society that should assume responsibility for developing an institutional framework for compensatory payments through a due procedure, and monitor the realization of the compensatory payments to the non-responsible victim.

With this principle in mind, an important fact about global warming that comes to the fore is that the economic activities of the present generation, which emit greenhouse gases and affect the personal identity of future generations, are subject to the exclusive control of the present generation. Thus, according to the principle of responsibility and compensation, the present generation should be held responsible for the problem of global warming in view of their full autonomy to control the policy choices to decide on a path of the future world.

Note that there exists a fundamental difference between the theoretical structure envisaged by Dworkin, Roemer and Fleurbaey and the structure of the problem of global warming, which prevents us from applying the Dworkin–Roemer–Fleurbaey theory directly to the problem of global warming. Recollect that the Dworkin–Roemer–Fleurbaey theory is implicitly assuming that either the subject who is responsible for the act of choice and the subject who is affected by the culmination outcome thereof are the same, or if they are not the same, they coexist at the same point of time. Consider the two previous examples again. The example of 'expensive tastes' is the case where the subjects are the same, whereas the example of 'accident by drunken driving' is the case where subjects differ, but the culprit and the victim coexist at the same point of time. As we have emphasized, however, one of the crucial features of the problem of global warming is that the responsible subjects are the past and present generations, whereas the seriously affected subjects are the distant future generations that do not exist at the time when the present generation makes its choice. Furthermore, depending on the culmination outcomes of the choice made by the present generation, the personal identity of the subjects who experience the consequences thereof is determined. It is in view of this unique feature of the problem of global warming that the meaning of choice of action made by the present generation with its concomitant responsibility requires more careful scrutiny.

Observe that the choice made by the present generation is nothing but the choice of a path from now to the indefinite future, which determines both the personal identity of the distant future generations as well as their living standard. From the viewpoint of the potential future generation, there is no way of going back along the historical path and making a choice of their own afresh. In this sense, the choice of the present generation is one-sided, external, and irrevocable. Thus, the primary meaning of responsibility that the present generation should hold for the future generations is the responsibility to choose a historical path that is accountable as a socially rationalizable choice based on some clear and rational standard of value judgement. This responsibility is worth the nomenclature of accountability.

16.6 A just future path

In the preceding section, we have established that the primary responsibility of the present generation is to choose a future path starting from the present that is accountable as a socially rationalizable choice based on some clear and rational principle. We have not, however, argued what principle should be adopted to evaluate various possible future paths. In other words, our conception of primary responsibility is independent of particular evaluation criterion of future paths.

We now take a particular evaluation criterion in the following analysis. To be consistent with the standpoint we have taken so far, we adopt the principle of responsibility and compensation itself for this purpose as well, and examine the implication of this principle for a future path.

To fix ideas, suppose that there are three alternative policies available for choice by the present generation. The first alternative is to abate global warming. The second alternative is to leave global warming unhindered, but improve the level of social capital accumulation. The third alternative is to leave global warming unhindered, and take no measure to improve the future social capital. Depending on which policy out of these alternatives is chosen by the present generation, the identity of individuals who would exist in the distant future would be determined, and their living standard would also be determined. Regardless of which individuals would exist in the future, the distant future generations cannot be held responsible for the choice of action by the present generation. In other words, the choice of action by the present generation is a non-responsible factor that determines the fate of the distant future generations from outside. If the present generation chooses the third policy alternative, the distant future people who then emerge would exist in harsh natural environments due to global warming, which is a consequence of a non-responsible choice made in the past, while receiving no compensating benefit from the accumulated social capital. In contrast, if the present generation chooses the second policy alternative, the distant future people, who would then emerge, would also be seriously affected by global warming, but they would be receiving compensating benefit by the medium of the accumulated social capital.

How should we evaluate these three policy alternatives? A most basic implication of the principle of responsibility and compensation is that the initial opportunities of individuals at the point of starting their lives, which are clearly determined by non-responsible factors for the individuals, should be equalized (Dworkin, 1981a,b).[2] In the context of global warming, the initial opportunities of future generations are determined by the actions of the present generation. In the above example, if the present generation chooses the third policy alternative, then the future generations who would then emerge would have very poor opportunities, compared to those of the present generation. Such inequality in the initial opportunities of different generations can never be judged as socially just according to the principle of responsibility and compensation. Hence, bearing our primary responsibility for the choice of future path, the present generation should not choose the third policy alternative because it cannot lead to a path that is made accountable as socially rational. In contrast, the choice of the first or the second policy alternative would not contradict our primary responsibility since we may make either path accountable as socially just based on the principle of responsibility and compensation.

Note that the improvement of social capital under the second policy alternative should not be construed as compensation for violating the right of the future generations to live in good natural environments. As was explained in section 16.4, the conception of such a right has a logical difficulty due to the non-identity problem. Accumulation of social capital is rather a means to realize a future path along which equality of the initial opportunities among generations is sustained.

16.7 Should we be single-handedly held responsible?

As we explained in section 16.2, not only the economic activities of the present generation but also those of the past generations since the Industrial

Revolution have contributed to an accumulation of greenhouse gases, causing global warming in the future. It is the past generations, not the present generation, who are responsible for past emissions of greenhouse gases. Hence, it might be argued that the present generation should bear responsibility to abate global warming only to the extent that they have contributed. We claim that this argument would contradict our conception of the primary responsibility of the present generation.

To fix ideas, suppose the following: (i) If we leave global warming unhindered, then the average temperature of the world will increase by three degrees in one hundred years; (ii) 50 per cent of this global warming is caused by the emissions of greenhouse gases of the present generation, but the remaining 50 per cent is due to the accumulation of greenhouse gases emitted by the past generations; (iii) If the temperature rises by 1.5 degrees, then it is very likely that some catastrophic disaster would occur; (iv) To avoid such catastrophe, the increase in the temperature must be at most 1 degree.

Now, assume that the present generation should bear responsibility to abate global warming only to the extent that they have contributed to it. Then, of the reduction of two degrees in the world temperature required to avoid catastrophe, the part that the present generation should be held responsible for would be only one degree reduction. The remaining one degree reduction should have been accomplished by the past generations. And if the present generation reduces the amount of greenhouse gas emissions that is only sufficient to decrease the world temperature by one degree, then some catastrophe will occur, and the initial opportunities of the future generations will become miserable. Such a future path can never be accountable as socially just according to the principle of responsibility and compensation. Hence, the choice of action of the present generation that leads to this path contradicts the primary responsibility of the present generation that we established in section 16.5. To accomplish the responsibility, therefore, the present generation must bear all the required reduction of two degrees in the world temperature.

It is true that the past generations should have reduced their emissions of greenhouse gases. But they are gone, and we cannot have them share the cost to abate global warming. If, in order to realize a future path that is accountable as socially rational, we have to bear the cost that the past generations should have shared, then we have no option other than doing so. Note that our primary responsibility for the choice of a future path should be accepted irrespective of the actions of the past generations. Of course, the feasible choices for the present generation depend upon the actions of the past generations. Given any historical path, however, the present generation should take the action to realize the future path starting from the present that is accountable as socially rational among all the feasible paths.

16.8 Concluding remarks

Consider again the two questions that we posed in the Introduction to this chapter. The first question was: (a) *Should we care about the possible plight of future generations? Those who might be most seriously affected by global warming would emerge after we are all gone, and we might never be forced to confront their plight.* The second question was: (b) *Should we be held single-handedly responsible? There are many past generations that should at least in part share the blame of what we are jointly responsible*. True enough, scientific evidence as well as public awareness of the problem of global warming is much more solid now than it used to be. But we need ethically appealing reasons for our duty to take actions unilaterally.

As an attempt to give a positive answer to the question (a), we formulated a modified version of the principle of responsibility and compensation in the form of responsibility for selecting a future path. We have also argued that the same principle can provide a reason why the present generation should assume sole responsibility to take action against global warming, even though past generations had not fulfilled their due responsibility. The modest purpose of this chapter would be served if we could initiate further discussions on problems (a) and (b) and their reasoned resolutions, which are logically prior to the problem of analytical description of rational choice of actions against global warming as well as the design of cost-sharing formulas.

Notes

1. In fact, it may be too much to say that there never exists a case where the Pareto principle can be applied to the issues related to global warming. Indeed, the Pareto principle can play a role in the situation where the policy a^{t^*} and the policy b^{t^*} have the same effect on the future generations, but the costs incurred to the present generation are different. The Pareto criterion applied to such a situation would function as the principle of cost minimization of policies that bring about the same culmination outcomes vis-à-vis global warming. However, in reality, there may rarely exist two different policies having the same consequences on the future generations with only the costs incurred to the present generation being different. For example, the policy choice between (i) whether the emission rights of greenhouse gases are traded or not, and (ii) if traded, in which way and under what restrictions the trade is to be implemented, would affect which areas on the earth the greenhouse gases would be reduced. This would then inevitably create some differences in the characteristics of the future generations. As is suggested by this example, the residual possibility of applying the Pareto criterion and/or the compensation principles to the problem of global warming seems to be quite narrowly circumscribed.
2. See, for instance, Vallentyne (2002) for the argument to the effect that justice requires that initial opportunities for advantage be equalized.

References

Coase, R. (1960) 'The Problem of Social Cost,' *Journal of Law and Economics*, vol. 3, pp. 1–44.

Dworkin, R. (1981a) 'What is Equality? Part 1: Equality of Welfare,' *Philosophy and Public Affairs*, vol. 10, pp. 185–246.

Dworkin, R. (1981b) 'What is Equality? Part 2: Equality of Resources,' *Philosophy and Public Affairs*, vol. 10, pp. 283–345.

Dworkin, R. (2000) *Sovereign Virtue: The Theory and Practice of Equality* (Cambridge, MA: Harvard University Press).

Fleurbaey, M. (1995) 'Equality and Responsibility,' *European Economic Review*, vol. 39, pp. 683–9.

Fleurbaey, M. (1998) 'Equality among Responsible Individuals,' in J.-J. Laslier et al. (eds), *Freedom in Economics: New Perspectives in Normative Analysis* (London: Routledge).

Mill, J. S. (1859) *On Liberty*. Reprinted in M. Warnock (ed.), *Utilitarianism* (London: Fontana, 1973).

Parfit, D. (1981) 'Future Generations, Further Problems,' *Philosophy and Public Affairs*, vol. 11, pp. 113–72.

Parfit, D. (1984) *Reasons and Persons* (Oxford: Oxford University Press).

Rawls, J. (1971) *A Theory of Justice* (Cambridge, MA: Harvard University Press), revised edition, 1999.

Roemer, J. E. (1985) 'Equity of Talent,' *Economics and Philosophy*, vol. 1, pp. 151–81.

Roemer, J. E. (1986) 'Equality of Resources Implies Equality of Welfare,' *Quarterly Journal of Economics*, vol. 101, pp. 751–84.

Vallentyne, P. (2002) 'Brute Luck, Option Luck, and Equality of Initial Opportunities,' *Ethics*, vol. 112, pp. 529–57.

17

Fundamental Incompatibility between Economic Efficiency, Intergenerational Equity and Sustainability*

Naoki Yoshihara

17.1 Introduction

In this chapter, we consider intergenerational resource allocations in production economies with joint production of public bads. The technological character in this production model is that the emission of the public bad is unavoidable in the production of private goods. Then, the emitted public bads are accumulated through every generation's economic activity, and the accumulated public bads stay in the stratosphere or under the ground over a long period of time. Thus, the current emission of public bads by the current generation may affect the living conditions of succeeding generations rather than itself. A typical example of such public bads is greenhouse gases (primarily carbon dioxide) in *global warming*.

Given this kind of technological structure, intergenerational resource allocations are discussed. There are possibly infinite streams of population, and each individual represents one generation. We assume that each generation exists on the earth in one particular period of time, and there is no structure of overlapping generations. Each generation is represented by an individual engaging in economic activity. So, he produces some private good by

* The author greatly appreciates Joaquim Silvestre's detailed comments and advice, which improved the paper toward its final version. An earlier version was presented at the IEA Roundtable Conference on Intergenerational Equity and Sustainability in March 2005. The author also thanks all the participants in the meeting, and in particular, Geir B. Asheim, Marc Fleurbaey, and Yongsheng Xu, for their useful comments. The author is also grateful to Tatsuyoshi Saijo and Koichi Tadenuma for their insightful discussions.

utilizing his labour as well as emitting a public bad. Moreover, he not only enjoys consumption of his leisure hours and the private good produced by himself, but should also be confronted with the negative effect of the accumulated public bads emitted by his ancestor generations. However, he is not affected by the current emission of the public bad which he produces. He can also invest some amount of the produced private good for *education*, which improves the technological knowledge of production utilized in the ages of his descendent generations.[1]

Thus, each *temporary allocation* by each generation consists of his consumption vector (a profile of his working hours, his consumption of private goods, and the accumulation of public bads confronted by him), his emission of public bads, and his investment of the private good for education. Then, an *intergenerationally feasible allocation* is defined by a historical sequence of each generation's temporary allocation, which is dynamically consistent with the technological conditions of production.

Our concern in this model is to examine the existence of *allocation rules*, which satisfy the three basic normative criteria, which are *economic efficiency*, *intergenerational equity*, and *environmental sustainability* respectively.[2] Economic efficiency is represented by the axiom of *Pareto efficiency*. In our model, every generation's preference is defined only on its own consumption space, so there is no altruistic aspect to each generation's rational choice behaviour. Thus, the definition of Pareto efficiency is a standard one except that the set of social alternatives is given by that of intergenerational allocations which is an infinite dimensional space. The criterion of environmental sustainability concerns improving the natural environment or leaving it as close to its initial condition as possible for the future generations.

As the axioms of intergenerational equity, we start by introducing the two traditional notions of equity: *equity as no-envy* (Foley, 1967) and *egalitarian equivalence* (Pazner and Schemeidler, 1978). Moreover, as weaker variants of the no-envy axiom, we define the axioms of *responsibility and compensation*, which were originally discussed and defined by Dworkin (1981) and Fleurbaey (1994, 1995) in the context of intragenerational resource allocations. Motivated by Suzumura and Tadenuma (2007), our axioms of responsibility and compensation represent the value judgements such that 'Any generations should be equally responsible for their descendent generations' living environments' and 'The generation who is more handicapped in their living environment should be compensated, being permitted to exploit the natural environment more intensively.' Both of them seem to be reasonable requirements in the context of intergenerational resource allocations with long-run negative externality.

It would be desirable to have an allocation rule satisfying the axioms, each of which respectively represents one of the above three basic normative criteria. Unfortunately, our answer to this question is essentially negative: Almost all of our main theorems discuss incompatibility between Pareto efficiency and any one of the axioms of intergenerational equity as well as the axiom of environmental sustainability.

In the following discussion, section 17.2 provides the basic model. Section 17.3 introduces the basic axioms of allocation rules in this model. Section 17.4 characterizes Pareto efficient allocations in this model, and section 17.5 discusses the main theorems. Finally, section 17.6 gives short concluding remarks.

17.2 The basic model

There is an infinite sequence of periods, $\mathbb{T} = \{1, 2, \ldots, t, \ldots\}$ with generic element t. Consider an economy at period t with one generation. To simplify the argument, we assume that every generation can live only in one period. Also, to focus on the intergenerational resource allocations, each generation is represented by one individual. Thus, the period t also implies '*generation t*.'

There is a (private) good $y \in \mathbb{R}_+$ which is produced from labour input $l \in \mathbb{R}_+$. The production process also involves emission of one public bad $z \in \mathbb{R}_+$. Thus, the production technology is represented by a function $f : \mathbb{R}^2_+ \to \mathbb{R}_+$ which is defined as: for any labour input $l \in \mathbb{R}_+$ and any public bad emissions $z \in \mathbb{R}_+$, $f(l, z) = y$. This function is assumed to be continuous, strictly increasing, concave, and $f(0, z) = 0$ & $f(l, 0) = 0$ for any $z \in \mathbb{R}_+$ and any $l \in \mathbb{R}_+$.

We impose an additional assumption on the technological progress in the production process. The function f is rewritten as: there exists a continuous, strictly increasing, and concave function $g : \mathbb{R}^2_+ \to \mathbb{R}_+$ and a positive number $h \in \mathbb{R}_{++}$ such that for any $l \in \mathbb{R}_+$ and any $z \in \mathbb{R}_+$, $f(l, z) = g(h \cdot l, z)$. Note that the number h indicates a level of *technological knowledge* or '*human capital*'. We consider that the value h is valuable in each generation. Thus, the production technology at generation t is denoted by: $f^t(l, z) = g(h^t \cdot l, z)$. The value h^t is determined by the combination of the knowledge h^{t-1} for any $t \in \mathbb{T}\backslash\{1\}$, which is accumulated from first generation to $(t-1)$-th generation, and the investment for education $w^{t-1} \in \mathbb{R}_+$ by the generation $t-1$. For $t = 1$, $h^1 = h^0 \in \mathbb{R}_{++}$, where h^0 is the level of human capital given at the age of pre-history. Thus, there is an *increasing* and *continuous* function H such that $h^t = H(h^{t-1}, w^{t-1})$ for any $t \in \mathbb{T}\backslash\{1\}$. The investment w^{t-1} is financed from the private good produced by the generation $t-1$. Suppose in the following discussion, that there is an *upper bound* $\bar{z}$ of the public bad emission. Also, suppose that $h^t = h^{t-1} = H(h^{t-1}, 0)$ for any $t \in \mathbb{T}\backslash\{1\}$. Sometimes, we discuss the case that H is constant. Note that the function H is *constant* if for any $h \in \mathbb{R}_{++}$ and any $w \in \mathbb{R}_+$, $h = H(h, w)$.

Each and every generation $t \in \mathbb{T}$ has the common consumption space $X \equiv [0, \bar{l}] \times \mathbb{R}_+ \times \mathbb{R}_+$, where $0 < \bar{l} < +\infty$ is the common upper bound of labour-leisure time, with the following generic formulas of consumption vector: $x^t = (l^t, y^t, Z^{t-1})$. The first two components of the vector x_t indicates that the person in the generation t supplies l_t amount of labour hour, and consumes y_t

amount of private good produced at period t, which is a standard argument. In contrast, the last component needs additional explanation: Z^{t-1} indicates the accumulated amount of public bads emitted at each period from $1 \in \mathbb{T}$ until $t-1 \in \mathbb{T}$. That is, if $z^{t'}$ is the amount of public bads emitted by the generation $t' = 1, \ldots, t-1$, then the consumption level of public bads by the generation t is defined by

$$Z^{t-1} \equiv \delta^{t-1} \cdot Z^0 + \sum_{t'=1}^{t-1} \delta^{(t-1)-t'} \cdot z^{t'}$$

where $\delta \in (0,1)$ is the natural rate of depletion at each period, and Z^0 is the *initial endowment of public bads* which came previously. Note that every generation t does not suffer from the current emission of the public bad z^t. It suffers from the accumulated amount of the public bad emitted by their ancestors up to the previous generation. This is the stylized fact of public bads consumption such as the problem of global warming.

Each and every generation $t \in \mathbb{T}$ is characterized by preference relation R_t on X. This relation is complete and transitive on X. For any $x^t, \widetilde{x}^t \in X$, $(x^t, \widetilde{x}^t) \in R_t$ means that x^t is at least as good as $\widetilde{x}^t$ according to t's preference. $P(R_t)$ and $I(R_t)$ denote, respectively, the strict preference relation and the indifference relation corresponding to R_t, viz., $(x^t, \widetilde{x}^t) \in P(R_t)$ if and only if $[(x^t, \widetilde{x}^t) \in R_t$ & $(\widetilde{x}^t, x^t) \notin R_t]$, and $(x^t, \widetilde{x}^t) \in I(R_t)$ if and only if $[(x^t, \widetilde{x}^t) \in R_t$ & $(\widetilde{x}^t, x^t) \in R_t]$. Also, R_t is assumed to be *strictly monotonic* (decreasing in labour time and public bad, and increasing in the share of output) on $(0, \bar{l}) \times \mathbb{R}_{++} \times \mathbb{R}_+$,[3] *continuous* and *convex* on X. The universal class of such preference relations is denoted by $\mathcal{R}$.

Given a stock of the public bad Z^{t-1}, an accumulated knowledge h^{t-1}, and an investment w^{t-1}, a pair $\mathbf{a}^t = (x^t, z^t, w^t) \in X \times \mathbb{R}_+ \times \mathbb{R}_+$ is a *temporarily feasible allocation for generation t* in $(Z^{t-1}, h^{t-1}, w^{t-1})$ if $x^t = (l^t, y^t, Z^{t-1})$, $h^t = H(h^{t-1}, w^{t-1})$, and $g(h^t \cdot l^t, z^t) \geq y^t + w^t$. The set of temporarily feasible allocations for generation t in $(Z^{t-1}, h^{t-1}, w^{t-1})$ is denoted by $A^t(Z^{t-1}, h^{t-1}, w^{t-1})$.

Given an initial stock of the public bad Z^0 and an initial endowment of human capital h^0, a historical sequence $\mathbf{a} = (\mathbf{a}^t)_{t \in \mathbb{T}}, = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in (X \times \mathbb{R}_+ \times \mathbb{R}_+)^\infty$ is an (*intergenerationally*) *feasible allocation in* (Z^0, h^0) if for all $t \in \mathbb{T}$, the stock of the public bad in this period is characterized by $Z^{t-1} \equiv \delta^{t-1} \cdot Z^0 + \sum_{t'=1}^{t-1} \delta^{(t-1)-t'} \cdot z^{t'}$, and $\mathbf{a}^t$ is the temporarily feasible allocation for generation t in $(Z^{t-1}, h^{t-1}, w^{t-1})$ at the period t, and the component of the public bad consumption of x^t is Z^{t-1}. Fixing (Z^0, h^0) in the following discussion, the set of feasible allocations is denoted by A.

Let us define an *economy* by a historical sequence of all generations' preferences $\mathbf{R}^{\mathbb{T}} = (R_t)_{t \in \mathbb{T}} \in \mathcal{R}^\infty$. An *allocation rule* is a correspondence φ: $\mathcal{R}^\infty \twoheadrightarrow A$ which associates to each $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^\infty$, a non-empty subset $\varphi(\mathbf{R}^{\mathbb{T}})$ of A.

17.3 Basic axioms for allocation rules

In this section, we introduce basic axioms, each of which represents an indispensable value in the problem of intergenerational resource allocations under economies with the long-run influence of negative externality. Those axioms are categorized into the three normative viewpoints: *economic efficiency*, *intergenerational equity* and *environmental sustainability*.

17.3.1 Axioms of economic efficiency

First, we discuss axioms of economic efficiency. Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, for each $t \in \mathbb{T}$ and each $(Z^{t-1}, h^{t-1}, w^{t-1})$, a temporarily feasible allocation $\mathbf{a}^t = (x^t, z^t, w^t) \in A^t(Z^{t-1}, h^{t-1}, w^{t-1})$ for generation t is *temporary Pareto efficient at* $\mathbf{R}^{\mathbb{T}}$ if there is no temporarily feasible allocation $\widetilde{\mathbf{a}}^t = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t) \in A^t(Z^{t-1}, h^{t-1}, w^{t-1})$ for generation t such that $(\widetilde{x}^t, x^t) \in P(R_t)$. Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, a feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t \in \mathbb{T}} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in A$ is *Pareto efficient at* $\mathbf{R}^{\mathbb{T}}$ if there is no feasible allocation $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t \in \mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)_{t \in \mathbb{T}} \in A$ such that for any $t \in \mathbb{T}$, $(\widetilde{x}^t, x^t) \in R_t$ holds, and there exists $t \in \mathbb{T}$ such that $(\widetilde{x}^t, x^t) \in P(R_t)$.

Now, we are ready to introduce a well-known axiom on allocation rules:

Pareto Efficiency (PE): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} \in \varphi(\mathbf{R}^{\mathbb{T}})$, $\mathbf{a}$ *is* Pareto efficient at $\mathbf{R}^{\mathbb{T}}$.

Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, for each $t \in \mathbb{T}$ and each $(Z^{t-1}, h^{t-1}, w^{t-1})$, a temporarily feasible allocation $\mathbf{a}^t = (x^t, z^t, w^t) \in A^t(Z^{t-1}, h^{t-1}, w^{t-1})$ for generation t is *temporarily selfish* if $(z^t, w^t) = (\overline{z}, 0)$ and there is no other $\widetilde{\mathbf{a}}^t = (\widetilde{x}^t, \overline{z}, 0) \in A^t(Z^{t-1}, h^{t-1}, w^{t-1})$ such that $(\widetilde{x}^t, x^t) \in P(R_t)$. Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, a feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t \in \mathbb{T}} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in A$ is *selfish* if for each $t \in \mathbb{T}$, $(z^t, w^t) = (\overline{z}, 0)$ and there is no other $\widetilde{\mathbf{a}}^t = (\widetilde{x}^t, \overline{z}, 0) \in A^t(Z^{t-1}, h^{t-1}, 0)$ such that $(\widetilde{x}^t, x^t) \in P(R_t)$.

Lemma 1: *Given an economy* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, *a* selfish allocation $\mathbf{a} = (\mathbf{a}^t)_{t \in \mathbb{T}} = (x^t, \overline{z}, 0)_{t \in \mathbb{T}} \in A$ *is* Pareto efficient at $\mathbf{R}^{\mathbb{T}}$.

Proof: Consider $t = 1$. Clearly, there is no other temporarily feasible allocation $\widetilde{\mathbf{a}}^1 = (\widetilde{x}^1, \widetilde{z}^1, \widetilde{w}^1) \in A^1(Z^0, h^0, 0)$ such that $(\widetilde{x}^1, x^1) \in P(R_1)$. Suppose $\widetilde{\mathbf{a}}^1 \in A^1(Z^0, h^0, 0)$ such that $(\widetilde{x}^1, x^1) \in I(R_1)$. Since $z^1 = \overline{z}$, it should be $\widetilde{z}^1 = \overline{z}$. Then, any feasible allocation $\widetilde{\mathbf{a}} \in A$ with $\widetilde{\mathbf{a}}^1$ at the first period should be unable to have $\widetilde{\mathbf{a}}^2 = (\widetilde{x}^2, \widetilde{z}^2, \widetilde{w}^2) \in A^2(Z^0, h^0, 0)$ such that $(\widetilde{x}^2, x^2) \in P(R_2)$. By repeating this procedure infinitely, we can see that there is no other feasible allocation which Pareto-dominates $\mathbf{a}$. □

Although **PE** is a fundamental requirement of economic efficiency, **Lemma 1** indicates that the set of Pareto efficient allocations contains selfish allocations in this model. Note that implementation of selfish allocations cannot resolve the issue of negative externality. In particular, if the negative externality by the emissions of public bads leads the society to the crisis of human sustainability, the implementation of selfish allocations indicates an undesirable situation. Thus, we would like to require:

Non-Selfish Pareto Efficiency (NSPE): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} \in \varphi(\mathbf{R}^{\mathbb{T}})$, $\mathbf{a}$ *is* non-selfish Pareto efficient at $\mathbf{R}^{\mathbb{T}}$.

Thus, the axiom **NSPE** requires not only efficient allocations of resources among generations, but also implementation of some policies for regulating the public bads emissions and of educational investments for future generations.

We will also discuss a second-best notion of efficiency. This notion requires a constrained efficient allocation of resources by some policies of regulation and investment. Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, for each $t \in \mathbb{T}$ and each $(Z^{t-1}, h^{t-1}, w^{t-1})$, a temporarily feasible allocation $\mathbf{a}^t = (x^t, z^t, w^t) \in A^t(Z^{t-1}, h^{t-1}, w^{t-1})$ for generation t is (z^t, w^t)-*constrained temporary Pareto efficient at* $\mathbf{R}^{\mathbb{T}}$ if there is no temporarily feasible allocation $\widetilde{\mathbf{a}}^t = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t) \in A^t(Z^{t-1}, h^{t-1}, w^{t-1})$ for generation t such that $(\widetilde{z}^t, \widetilde{w}^t) = (z^t, w^t)$ and $(\widetilde{x}^t, x^t) \in P(R_t)$. It is clear that if $\mathbf{a} = (\mathbf{a}^t)_{t \in \mathbb{T}} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in A$ is Pareto efficient at $\mathbf{R}^{\mathbb{T}}$, then for any $t \in \mathbb{T}$, $\mathbf{a}^t = (x^t, z^t, w^t)$ is (z^t, w^t)-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$.

Given a feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t \in \mathbb{T}} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in A$, $(z^t, w^t)_{t \in \mathbb{T}}$ is a *feasible sequence of public bads provisions and investments for new technological knowledge*. Given such a feasible sequence $(z^t, w^t)_{t \in \mathbb{T}}$, let $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t \in \mathbb{T}} = (\widetilde{x}^t, z^t, w^t)_{t \in \mathbb{T}} \in A$ be another feasible allocation whose components of public bads and investments are $(z^t, w^t)_{t \in \mathbb{T}}$. Let us denote the set of such feasible allocations by $A(z^t, w^t)_{t \in \mathbb{T}}$ when the feasible sequence $(z^t, w^t)_{t \in \mathbb{T}}$ is given. Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, and given a feasible sequence $(z^t, w^t)_{t \in \mathbb{T}}$, a feasible allocation $\mathbf{a}^* = (\mathbf{a}^{*t})_{t \in \mathbb{T}} = (x^{*t}, z^t, w^t)_{t \in \mathbb{T}} \in A$ is $(z^t, w^t)_{t \in \mathbb{T}}$-*constrained Pareto efficient at* $\mathbf{R}^{\mathbb{T}}$ if there is no other feasible allocation $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t \in \mathbb{T}}, = (\widetilde{x}^t, z^t, w^t)_{t \in \mathbb{T}} \in A(z^t, w^t)_{t \in \mathbb{T}}$ such that for any $t \in \mathbb{T}$, $(\widetilde{x}^t, x^t) \in R_t$ holds, and there exists $t' \in \mathbb{T}$ such that $(\widetilde{x}^{t'}, x^{t'}) \in P(R_{t'})$. Note that for any feasible allocation $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in A$, it is $(z^t, w^t)_{t \in \mathbb{T}}$-constrained Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ if and only if for any $t \in \mathbb{T}$, $\mathbf{a}^t = (x^t, z^t, w^t)$ is (z^t, w^t)-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$.

17.3.2 Axioms on intergenerational equity

Here we discuss axioms of intergenerational equity. In the first place, the following axiom is relevant to *equity* in terms of subjective well-being. This is an extension of the no-envy principle (Foley, 1967) to the problem of intergenerational resource allocations:

No-Envy (NE): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$, *we have that: for any* $t \in \mathbb{T}$, $(x^t, x^{t'}) \in R_t$ *holds for any* $t' \in \mathbb{T}$.

The following five axioms are weaker versions of the no-envy axiom:

Equal Welfare for Equal Preferences (EWEP): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$, *we have that: for any* $t, t' \in \mathbb{T}$, *if* $R_t = R_{t'}$, *then one of the following three holds*: $(x^t, x^{t'}) \in I(R_t)$, $(x^t, x^{t'}) \in R_t$ *for* $x^t = (l^t, 0, Z^{t-1})$, *and* $(x^{t'}, x^t) \in R_t$ *for* $x^{t'} = (l^{t'}, 0, Z^{t'-1})$.

Equal Welfare for Uniform Preferences (EWUP): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$, *we have that: if for any* $t, t' \in \mathbb{T}$, $R_t = R_{t'}$, *then for any* $t, t' \in \mathbb{T}$, *one of the following three holds*: $(x^t, x^{t'}) \in I(R_t)$, $(x^t, x^{t'}) \in R_t$ *for* $x^t = (l^t, 0, Z^{t-1})$, *and* $(x^{t'}, x^t) \in R_t$ *for* $x^{t'} = (l^{t'}, 0, Z^{t'-1})$.

The above two axioms originated with Fleurbaey (1994, 1995), who discussed intragenerational resource allocations under pure exchange economies. These are axioms of *compensation for "more handicapped generations*." Note here the "more handicapped generations" means the generations endowed with more amount of accumulated public bads and/or less accumulation of human capital.

It is easy to see that **EWUP** is a weaker variant of **EWEP**. The next axiom is another weaker variant of **EWEP**:

Undomination among Equal Preferences (UNEP): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$, *we have that: for any* $t, t' \in \mathbb{T}$, *if* $R_t = R_{t'}$, *then* $(-Z^{t-1}, h^t) > (-Z^{t'-1}, h^{t'}) \Rightarrow [z^t < z^{t'} \text{or } w^t > w^{t'}]$.[4]

This axiom says that the 'more handicapped generation' has a right to produce and utilize more resources for only his own consumption.

The next three axioms are of *responsibility* axioms. The first two are a variation of the 'No-envy among equal skills', which was originally discussed by Fleurbaey and Maniquet (1996) in the context of intragenerational resource allocations under production economies:

No-Envy among Equal-Endowed Generations (NEEG): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$, *we have that: for any* $t, t' \in \mathbb{T}$, *if* $Z^{t-1} = Z^{t'-1}$ *and* $h^t = h^{t'}$, *where* Z^{t-1} *(resp.* $Z^{t'-1}$*) is the third component of the consumption vector* x^t *(resp.* $x^{t'}$*), then* $(x^t, x^{t'}) \in R_t$ *and* $(x^{t'}, x^t) \in R_{t'}$ *hold.*

No-Envy among Uniform-Endowed Generations (NEUG): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$, *we have that: if for any* $t, t' \in \mathbb{T}$, $Z^{t-1} = Z^{t'-1}$ *and* $h^t = h^{t'}$ *, where* Z^{t-1} *(resp.* $Z^{t'-1}$*) is the third component of the consumption vector* x^t *(resp.* $x^{t'}$*), then for any* $t, t' \in \mathbb{T}$ *,* $(x^t, x^{t'}) \in R_t$ *and* $(x^{t'}, x^t) \in R_{t'}$ *hold.*

The third axiom of responsibility is introduced as follows:

Responsibility for Future Generations (RFG): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$, *we have that: for any* $t, t' \in \mathbb{T}$, *if* $Z^{t-1} = Z^{t'-1}$ *and* $h^t = h^{t'}$, *where* Z^{t-1} *(resp.* $Z^{t'-1}$*) is the third component of the consumption vector* x^t *(resp.* $x^{t'}$*), then* $[z^t > z^{t'}] \Rightarrow [w^t > w^{t'}]$.

This axiom requires responsibility of the current generation to keep the 'living environment' as well as possible for future generations. Note that **NEEG** implies **RFG**.

Although the above axioms are of equity as no-envy and its weaker variations, we can also discuss a variation of the egalitarian-equivalent principle (Pazner and Schmeidler, 1978) in this problem of intergenerational resource allocations. When we discuss the egalitarian-equivalence axiom here, let us assume that $(1-\delta)Z^0 \leq \overline{z}$. Let $z^* \leq \overline{z}$ be a social reference level of public bads emission. Then, let us define (Z^0, h^0, z^*) as a *social reference level of 'natural environments'*. Given any generation $t \in \mathbb{T}$ with his preference R_t, let $x^t(R_t; Z^0, h^0, z^*)$ be t's ideal consumption vector which is maximal with respect to R_t whenever he is faced with production condition $g(h^0 l, z^*)$ and the stock of public bads Z^0. Now, we are ready to discuss a variation of the egalitarian-equivalent principle in this context:

z^***-Egalitarian Equivalence** (z^***-EE**): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$*, we have that: for any* $t \in \mathbb{T}$*,* $(x^t, x^t(R_t; Z^0, h^0, z^*)) \in I(R_t)$ *holds.*

This axiom is a requirement of equal opportunity for welfare among generations. In particular, if z^* is given by $z^* = (1-\delta)Z^0$, then the social reference profile (Z^0, h^0, z^*) indicates that every generation is guaranteed an initial natural environment (Z^0, h^0) by the preceding generations, and they also guarantee this environment for a descendant generation by restricting the public bads emission to z^*. Thus, the axiom z^*-**EE** guarantees every generation equally the welfare level which is maximal under the constraint (Z^0, h^0, z^*).

17.3.3 Axioms of environmental sustainability

We can also consider other axioms to judge the wellness of intergenerational resource allocations, which are relevant to *sustainability*. By sustainability, we may consider at least two meanings in the environmental literature. One is of the 'natural environmentalist' who insists on the intrinsic value of natural environments, where such a value is not necessarily relevant to the welfare of human beings. So, sustainability should mean for the 'natural environmentalist,' that the historical sequence of stocks of public bads is non-increasing in $\mathbb{T}$.

As the axiom of sustainability for the 'natural environmentalist', we define the followings:

Public Bads Monotonicity (PBM): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} = ((l^t, y^t, Z^{t-1}), z^t, w^t)_{t \in \mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$*, it holds that* $Z^{t'-1} \geq Z^{t''-1}$ *whenever* $t' \leq t''$*.*

This axiom is well-defined. Take a historical sequence of public bad provision $\mathbf{z} = (z^t)_{t \in \mathbb{T}}$ such that $z^1 \leq (1-\delta) \cdot Z^0$, and for any other $t \in \mathbb{T} \backslash \{1\}$,

$z^t = (1-\delta)\cdot[\delta Z^0 + z^1]$. Then, the emission of z^1 is compatible with temporarily feasible allocations, and for any $t \in \mathbb{T}$, $Z^{t-1} = Z^1 = \delta Z^0 + z^1 \leq Z^0$ holds. These facts imply that it is possible to construct an allocation rule which satisfies the axiom **PBM**.

The other meaning is from the 'humanist' standpoint. For the 'humanist' sustainability means that the historical sequence of *human welfare* is non-increasing in $\mathbb{T}$.[5] In this approach, an important issue is how to measure human welfare. Since each generation's preference is ordinally measurable and intergenerationally noncomparable in this model, we cannot use it as for measuring each generation's welfare: the requisite of the non-increasing of human welfare over periods implicitly assumes the existence of intergenerationally comparable welfare units.

In this chapter, we assume the existence of an objective welfare measure. A typical example of such a measure can be found in the theory of *functionings and capability* developed by Sen (1980, 1985): the welfare measure is a *representation of some ordering relation defined over alternative capabilities* that human beings can enjoy. This is formulated by an ordering J defined over $X \times \mathbb{R}_+$, where this $\mathbb{R}_+$ is the space for human capital. Thus, for any $(x, h), (x', h') \in X \times \mathbb{R}_+$, $((x, h), (x', h')) \in J$ implies that having the consumption vector x and the knowledge h is at least as desirable for human beings as having the consumption vector x' and the knowledge h'. Then:

J-Reference Human Development (J-HD): *For all* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ *and* $\mathbf{a} = (x^t, z^t, w^t)_{t\in\mathbb{T}} = ((l^t, y^t, Z^{t-1}), z^t, w^t)_{t\in\mathbb{T}} \in \varphi(\mathbf{R}^{\mathbb{T}})$, *it holds that* $((x^{t''}, h^{t''}), (x^{t'}, h^{t'})) \in J$ *whenever* $t' \leq t''$.

Note that the meaning of human development is based upon the property of the ordering J. In the following discussion, we naturally assume that J is *continuous* and *strictly monotonic in* $X \times \mathbb{R}_+$ (decreasing in labour hours and public bads, and increasing in the share of output and level of knowledge), and *convex on* $X \times \{h\}$ for any $h \in \mathbb{R}_+$. Thus, for every generation, an inherited higher level of knowledge and a lower level of the stock of public bads can enhance their objective welfare, while bequeathing a lower level of educational investment and a higher level of public bads emission can make their descendent generations worse off in terms of the objective welfare measure J.

17.4 Characterizations of intergenerational Pareto efficiency

Before examining the possibility of allocation rules satisfying the axioms relevant to economic efficiency, intergenerational equity, and sustainability, we characterize Pareto efficient allocations in this model. The following lemma gives us a necessary and sufficient condition for Pareto efficiency.

Lemma 2: *Given an economy* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$*, a feasible allocation* $\mathbf{a} = (\mathbf{a}^t)_{t\in\mathbb{T}} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ *is* Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ *if and only if the following condition holds*:
($\star$) *For any* $\bar{t} \in \mathbb{T}$ *and any* $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t\in\mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}} \in A$ *with* $(x^t, \widetilde{x}^t) \in I(R_t)$ *for any* $t < \bar{t}$*, if* $(\widetilde{x}^{\bar{t}}, x^{\bar{t}}) \in P(R_{\bar{t}})$*, then there exists* $\bar{t}' \in \mathbb{T}$ *such that* $\bar{t}' > \bar{t}$ *and* $(x^{\bar{t}'}, \widetilde{x}^{\bar{t}'}) \in P(R_{\bar{t}'})$.

This lemma is almost the definition of Pareto efficiency, so we omit proof. We can also have the necessary and sufficient condition for Pareto efficiency, in the specific case of the constant H.

Lemma 3: *Assume* H *is constant. Given an economy* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$*, a feasible allocation* $\mathbf{a} = (\mathbf{a}^t)_{t\in\mathbb{T}} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ *is* Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ *if and only if the following condition holds*:
(*) *For any* $\bar{t} \in \mathbb{T}$ *and any* $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t\in\mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}} \in A$ *with* $(x^t, z^t, w^t) = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)$ *for any* $t < \bar{t}$*, if* $(\widetilde{x}^{\bar{t}}, x^{\bar{t}}) \in P(R_{\bar{t}})$*, then there exists* $\bar{t}' \in \mathbb{T}$ *such that* $\bar{t}' > \bar{t}$ *and* $(x^{\bar{t}'}, \widetilde{x}^{\bar{t}'}) \in P(R_{\bar{t}'})$.

Proof: Let us examine if $\mathbf{a}$ is Pareto efficient at $\mathbf{R}^{\mathbb{T}}$, then $\mathbf{a}$ meets the condition (*). Suppose $\mathbf{a}, = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ violates (*). Then, there exists $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t\in\mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}} \in A$ with some $\bar{t} \in \mathbb{T}$ such that $(x^t, z^t, w^t) = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)$ for any $t < \bar{t}$, $(\widetilde{x}^{\bar{t}}, x^{\bar{t}}) \in P(R_{\bar{t}})$, and $(\widetilde{x}^t, x^t) \in R_t$ for any $t > \bar{t}$. This implies $\mathbf{a}$ is Pareto-dominated by $\widetilde{\mathbf{a}}$.

Consider the inverse relation. Suppose $\mathbf{a} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ is not Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. Then, there is a feasible allocation $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t\in\mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}} \in A$ which Pareto-dominates $\mathbf{a}$ at $\mathbf{R}^{\mathbb{T}}$. Suppose there exists $\bar{t} \in \mathbb{T}\backslash\{1\}$ such that $(x^t, \widetilde{x}^t) \in I(R_t)$ for any $t < \bar{t}$, $(\widetilde{x}^{\bar{t}}, x^{\bar{t}}) \in P(R_{\bar{t}})$, and $(\widetilde{x}^t, x^t) \in R_t$ for any $t > \bar{t}$.

First, consider the case that for any $t < \bar{t}$, (x^t, z^t, w^t) is (z^t, w^t)-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. Then, for $t = 1$, $(\widetilde{x}^1, \widetilde{z}^1, \widetilde{w}^1) \neq (x^1, z^1, w^1)$ and $(x^1, \widetilde{x}^1) \in I(R_1)$ imply either $w^1 > 0$ or $\widetilde{z}^1 > z^1$. Since H is constant, if $w^1 > 0$, then the new allocation $\widehat{\mathbf{a}} \in A$ such that $\widehat{\mathbf{a}}^1 = ((l^1, y^1 + w^1, Z^0), z^t, 0^1)$ and $\widehat{\mathbf{a}}^t = (x^t, z^t, w^t)$ for any $t \in \mathbb{T}\backslash\{1\}$ has the property to violate the condition (*) for $\mathbf{a}$ at $\mathbf{R}^{\mathbb{T}}$. If $\widetilde{z}^1 > z^1$, then $(x^1, \widetilde{x}^1) \in I(R_1)$ implies that $(\widetilde{x}^1, \widetilde{z}^1, \widetilde{w}^1)$ is not $(\widetilde{z}^1, \widetilde{w}^1)$-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. Thus, the new allocation $\widehat{\mathbf{a}} \in A$ such that $\widehat{\mathbf{a}}^1 = (\widetilde{x}^1, \widetilde{z}^1, \widetilde{w}^1)$, where $\widehat{\mathbf{a}}^1$ is $(\widetilde{z}^1, \widetilde{w}^1)$-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$, and $\widehat{\mathbf{a}}^t = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)$ for any $t \in \mathbb{T}\backslash\{1\}$ has the property to violate the condition (*) for $\mathbf{a}$ at $\mathbf{R}^{\mathbb{T}}$.

If $(\widetilde{x}^1, \widetilde{z}^1, \widetilde{w}^1) = (x^1, z^1, w^1)$ and $(\widetilde{x}^2, \widetilde{z}^2, \widetilde{w}^2) \neq (x^2, z^2, w^2)$, then by a similar discussion to that for $(\widetilde{x}^1, \widetilde{z}^1, \widetilde{w}^1) \neq (x^1, z^1, w^1)$ in the previous paragraph, we can construct a new allocation $\widehat{\mathbf{a}} \in A$ which Pareto-dominates $\mathbf{a}$ at $\mathbf{R}^{\mathbb{T}}$ such that $\widehat{\mathbf{a}}^1 = (x^1, z^1, w^1)$ and $\widehat{\mathbf{a}}^2$ with $(\widehat{x}^2, x^2) \in P(R_2)$. This $\widehat{\mathbf{a}} \in A$ has the property to violate the condition (*) for $\mathbf{a}$ at $\mathbf{R}^{\mathbb{T}}$. In such a way, we can show that if $(x^t, \widetilde{x}^t) \in I(R_t)$ holds for any $t < \bar{t}$, then $(x^t, z^t, w^t) = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)$ holds for any $t < \bar{t}$. In this case, $\widetilde{\mathbf{a}}$ is the desired allocation which has the property to violate the condition (*) for $\mathbf{a}$ at $\mathbf{R}^{\mathbb{T}}$.

Secondly, consider the case that there exists $t' < \overline{t}$ such that $(x^{t'}, z^{t'}, w^{t'})$ is not $(z^{t'}, w^{t'})$-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. Then, it is easy to construct an alternative allocation $\widehat{\mathbf{a}} \in A$ such that $(\widehat{x}^{t'}, \widehat{z}^{t'}, \widehat{w}^{t'})$ is $(z^{t'}, w^{t'})$-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$, and $(\widehat{x}^{t}, \widehat{z}^{t}, \widehat{w}^{t}) = (x^t, z^t, w^t)$ holds for any other $t \neq t'$. Then, this allocation $\widehat{\mathbf{a}}$ violates the condition (*) for $\mathbf{a}$ at $\mathbf{R}^{\mathbb{T}}$. □

Let a feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t\in\mathbb{T}} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ be called a *no-investment allocation* if for any $t \in \mathbb{T}$, $w^t = 0$.

Lemma 4: *Assume H is constant. Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, let a feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t\in\mathbb{T}} = (x^t, z^t, 0)_{t\in\mathbb{T}} \in A$ be a no-investment allocation. Suppose for this allocation, there is any integer $k > 0$ such that*
(1) $z^t < \overline{z}$ for any $t < k$;
(2) $z^t < \overline{z}$ for any $t \in \mathbb{T}$ such that there exists a positive integer $n > 0$ with $nk < t < (n+1)k$; and
(3) $z^t = \overline{z}$ for any $t = nk$ and for any positive integer $n = 1, 2, \ldots,$.
*Moreover, for any $t \in \mathbb{T}$, $\mathbf{a}^t$ is $(z^t, 0)$-*constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. *Then,* $\mathbf{a}$ *is* Pareto efficient at $\mathbf{R}^{\mathbb{T}}$.

Proof: Suppose that there exists an alternative allocation $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t\in\mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}} \in A$ and a generation $\overline{t} \in \mathbb{T}$ such that $(x^t, z^t, w^t) = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)$ for any $t < \overline{t}$, and $(\widetilde{x}^{\overline{t}}, x^{\overline{t}}) \in P(R_{\overline{t}})$. Then, by definition of $\mathbf{a}$, $\widetilde{z}^{\overline{t}} > z^{\overline{t}}$ and $\widetilde{w}^{\overline{t}} = 0$. Note that there exists a positive integer $n > 0$ such that $(n-1)k < \overline{t} < nk$. Then, $z^{nk} = \overline{z}$. Suppose that for any $t \in \mathbb{T}$ with $\overline{t} < t < nk$, $(\widetilde{x}^t, x^t) \in R_t$. Then, $\widetilde{z}^{\overline{t}+1} > z^{\overline{t}+1}$ and $\widetilde{w}^{\overline{t}+1} = 0$, since $\widetilde{Z}^{\overline{t}} > Z^{\overline{t}}$. Thus, to keep $(\widetilde{x}^t, x^t) \in R_t$ for $t = \overline{t}+2$, it follows $\widetilde{z}^{\overline{t}+2} > z^{\overline{t}+2}$ and $\widetilde{w}^{\overline{t}+2} = 0$, since $\widetilde{Z}^{\overline{t}+1} > Z^{\overline{t}+1}$. In a similar way, $\widetilde{z}^t > z^t$ and $\widetilde{w}^t = 0$ for any $t \in \mathbb{T}$ with $\overline{t} < t < nk$. Thus, $\widetilde{Z}^{nk-1} > Z^{nk-1}$ and $z^{nk} = \overline{z}$ imply $(x^{nk}, \widetilde{x}^{nk}) \in P(R_{nk})$. By **Lemma 3**, $\mathbf{a}$ is Pareto efficient. □

The next lemma shows that if every generation is assigned a temporarily non-selfish allocation, such a feasible allocation cannot be Pareto efficient.

Lemma 5: *Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, let a feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t\in\mathbb{T}} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ be $(z^t, w^t)_{t\in\mathbb{T}}$-*constrained Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ *such that $z^t < \overline{z}$ for any $t \in \mathbb{T}$. Then,* $\mathbf{a}$ *is not* Pareto efficient at $\mathbf{R}^{\mathbb{T}}$.

Proof: Suppose that R_t is representable by a continuous real-valued function u^t. Consider an alternative allocation $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t\in\mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}} \in A$, which is $(\widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}}$-constrained Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ and is defined as follows:
(1) $\widetilde{w}^t = w^t$ for any $t \in \mathbb{T}$;
(2) $\widetilde{z}^t = z^t + \triangle z^t$ for $t = 1$, where $\triangle z^t > 0$ is small enough;
(3) $\widetilde{z}^t = z^t + \triangle z^t$ for $t \in \mathbb{T}\backslash\{1\}$ such that

$$\begin{aligned} &u^t(l^t, g(h^t l^t, z^t) - w^t, Z^{t-1}) - u^t(l^t, g(h^t l^t, z^t) - w^t, \widetilde{Z}^{t-1}) \\ &\leq u^t(l^t, g(h^t l^t, \widetilde{z}^t) - w^t, \widetilde{Z}^{t-1}) - u^t(l^t, g(h^t l^t, z^t) - w^t, \widetilde{Z}^{t-1}), \end{aligned}$$

where $\triangle Z^{t-1} = \widetilde{Z}^{t-1} - Z^{t-1} > 0$. We can find an appropriate $(\triangle z^t)_{t\in\mathbb{T}}$ which guarantees $z^t + \triangle z^t \leq \bar{z}$ for any $t \in \mathbb{T}$.

Let us show this. Given any small enough $\triangle z^1 > 0$, let $\triangle \mathbf{z}^{\mathbb{T}}(\triangle z^1) \in \mathbb{R}^{\infty}_{++}$ be a vector $(\triangle z^t)_{t\in\mathbb{T}}$ satisfying the above (2) and (3) with $\triangle z^1$ as its first component, such that for any other $(\triangle \widehat{z}^t)_{t\in\mathbb{T}}$ satisfying (2) and (3) with $\triangle \widehat{z}^1 = \triangle z^1$, $\triangle z^t \leq \triangle \widehat{z}^t$ holds for any $t \in \mathbb{T}$. By the continuity of u^t and g, the existence of such $\triangle \mathbf{z}^{\mathbb{T}}(\triangle z^1)$ is guaranteed for any small enough $\triangle z^1 > 0$. Note that each component of the vector $\triangle \mathbf{z}^{\mathbb{T}}(\triangle z^1)$ increases when $\triangle z^1$ increases. The mapping $\triangle \mathbf{z}^{\mathbb{T}}$ is also continuous at every small enough $\triangle z^1$.[6]

Suppose for some $\triangle \widehat{z}^1 > 0$, $\triangle \mathbf{z}^{\mathbb{T}}(\triangle \widehat{z}^1)$ has a subset $\mathbb{N}$ of $\mathbb{T}$ such that for any $t \in \mathbb{N}$, t has $\triangle z^t(\triangle \widehat{z}^1) > \bar{z} - z^t$ in this vector, where $\triangle z^t(\triangle \widehat{z}^1)$ means the t-th component of the vector $\triangle \mathbf{z}^{\mathbb{T}}(\triangle \widehat{z}^1)$. However, since $\triangle \mathbf{z}^{\mathbb{T}}(\triangle z^1) \to \mathbf{0} \in \mathbb{R}^{\infty}_{+}$ as $\triangle z^1 \to 0$, we find an appropriately small enough $\triangle z^{*1} > 0$ such that for any $t \in \mathbb{T}$, $\triangle z^t(\triangle z^{*1}) \leq \bar{z} - z^t$ by the increasing and continuous property of the mapping $\triangle \mathbf{z}^{\mathbb{T}}$.

By construction, $(\widetilde{x}^1, x^1) \in P(R_1)$. Moreover, $(\widetilde{x}^t, x^t) \in R_t$ for any $t \in \mathbb{T}\backslash\{1\}$. This implies $\widetilde{\mathbf{a}}$ Pareto dominates $\mathbf{a}$. □

Lemma 5 deserves some comment. It shows that if generations fail to emit the maximal public bads, then a Pareto improving allocation can be constructed by increasing the sequence of public bads $\{\widetilde{z}^t\}_{t\in\mathbb{T}}$ in such a way that each generation compensates the increased inherited public bads, in turn by appropriately increasing the amount of public bads bequeathed. This type of situation would not work in a finite-horizon economy.

Throughout the above arguments on Pareto efficiency in this intergenerational resource allocations, we can summarize as in the following:

Proposition 1: *Given an economy* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, *let a feasible allocation* $\mathbf{a} = (\mathbf{a}^t)_{t\in\mathbb{T}} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ *be* $(z^t, w^t)_{t\in\mathbb{T}}$-constrained Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. *Then,* $\mathbf{a}$ *is* Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ *only if for any* $t \in \mathbb{T}$, $[z^t < \bar{z}]$ *implies* $[\exists t' \in \mathbb{T}$ s.t. $t' > t$, $z^{t'} = \bar{z}$ and $w^{t'} = 0]$.

Proof: Suppose that there exists $\bar{t} \in \mathbb{T}$ such that $z^{\bar{t}} = \bar{z}$ holds, and for any $t' \in \mathbb{T}$ with $t' > \bar{t}$, $z^{t'} < \bar{z}$ or $w^{t'} > 0$ holds. Then, consider another feasible allocation $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t\in\mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}} \in A$, which is $(\widetilde{z}^t, \widetilde{w}^t)_{t\in\mathbb{T}}$-constrained Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ and is defined as follows:

(1) for any $t < \bar{t}$, $\widetilde{\mathbf{a}}^t = \mathbf{a}^t$;

(2) for $t = \bar{t}$, $\widetilde{z}^t = z^t + \triangle z^t$ & $\widetilde{w}^t = w^t$, where $\triangle z^t > 0$ is small enough;

(3) for $t \in \mathbb{T}\backslash\{\bar{t}\}$ with $t > \bar{t}$, either (i) $\widetilde{z}^t = z^t + \triangle z^t$ & $\widetilde{w}^t = w^t$, or (ii) $\widetilde{z}^t \geq z^t$ & $\widetilde{w}^t = w^t - \triangle w^t$ if $w^t > 0$, such that

$$
\begin{aligned}
&u^t(l^t, g(h^t l^t, z^t) - w^t, Z^{t-1}) - u^t(l^t, g(\widetilde{h}^t l^t, z^t) - w^t, \widetilde{Z}^{t-1}) \\
&\leq u^t(l^t, g(\widetilde{h}^t l^t, \widetilde{z}^t) - \widetilde{w}^t, \widetilde{Z}^{t-1}) - u^t(l^t, g(\widetilde{h}^t l^t, z^t) - w^t, \widetilde{Z}^{t-1}),
\end{aligned}
$$

where $\widetilde{Z}^{t-1} = Z^{t-1} + \triangle Z^{t-1}$ for $\triangle Z^{t-1} = \sum_{t'=1}^{t-1} \delta^{(t-1)-t'} \cdot \triangle z^{t'}$, $\widetilde{h}^{\bar{t}+1} = H(h^{\bar{t}}, \widetilde{w}^{\bar{t}}) = H(h^{\bar{t}}, w^{\bar{t}}) = h^{\bar{t}+1}$, and $\widetilde{h}^t = H(\widetilde{h}^{t-1}, \widetilde{w}^{t-1})$ for any $t \in \mathbb{T} \backslash \{\bar{t}+1\}$ with $t > \bar{t}$. Through an argument similar to that in the proof of Lemma 5, we can confirm that $\triangle z^t \leq \bar{z} - z^t$ and $\triangle w^t \in [0, w^t]$ for any $t \in \mathbb{T} \backslash \{\bar{t}+1\}$ with $t > \bar{t}$. By construction, $(\widetilde{x}^{\bar{t}+1}, x^{\bar{t}+1}) \in P(R_{\bar{t}+1})$. Moreover, $(\widetilde{x}^t, x^t) \in R_t$ for any $t \in \mathbb{T} \backslash \{\bar{t}+1\}$ with $t > \bar{t}$. This implies $\widetilde{\mathbf{a}}$ Pareto dominates $\mathbf{a}$. □

Thus, **Proposition 1** shows that every Pareto efficient allocation needs an infinite subset of generations who enjoy 'selfish' consumptions. Otherwise, there exists a Pareto improving allocation with an increased sequence of public bads emissions as discussed in the comment for **Lemma 5**.

Proposition 2: *Assume H is constant. Given an economy* $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, *let a feasible allocation* $\mathbf{a} = (\mathbf{a}^t)_{t \in \mathbb{T}} = (x^t, z^t, 0^t)_{t \in \mathbb{T}} \in A$ *be* $(z^t, 0^t)_{t \in \mathbb{T}}$-constrained Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. *Then,* $\mathbf{a}$ *is* Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ *if and only if the following condition holds:*

(*) *For any* $t \in \mathbb{T}$, $[z^t < \bar{z}]$ *implies* $[\exists t' \in \mathbb{T}$ s.t. $t' > t$, $z^{t'} = \bar{z}$ and $w^{t'} = 0]$.

Proof: Suppose that for any $t \in \mathbb{T}$, $[z^t < \bar{z}]$ implies

$$\left[\exists t' \in \mathbb{T} \text{ s.t. } t' > t,\, z^{t'} = \bar{z} \text{ and } w^{t'} = 0\right].$$

Suppose that there exists an alternative allocation $\widetilde{\mathbf{a}} = (\widetilde{\mathbf{a}}^t)_{t \in \mathbb{T}} = (\widetilde{x}^t, \widetilde{z}^t, 0^t)_{t \in \mathbb{T}} \in A$ and a generation $\bar{t} \in \mathbb{T}$ such that $(x^t, z^t, 0^t) = (\widetilde{x}^t, \widetilde{z}^t, 0^t)$ for any $t < \bar{t}$, and $(\widetilde{x}^{\bar{t}}, x^{\bar{t}}) \in P(R_{\bar{t}})$. Then, by definition of $\mathbf{a}$, $\widetilde{z}^{\bar{t}} > z^{\bar{t}}$. Note there exists $t' \in \mathbb{T}$ such that $t' > \bar{t}$, $z^{t'} = \bar{z}$ and $w^{t'} = 0$. Then, following the proof of **Lemma 4**, we can see that if $(\widetilde{x}^t, x^t) \in R_t$ for any $t \in \mathbb{T}$ with $\bar{t} < t < t'$, then $(x^{t'}, \widetilde{x}^{t'}) \in P(R_{t'})$. Thus, by **Lemma 3**, $\mathbf{a}$ is Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. □

The above two characterization results indicate that any regulation policy for the public bad emissions is incompatible with Pareto efficiency, whenever it requires $z^t < \bar{z}$ for any $t \in \mathbb{T}$. However, it seems to be reasonable, from the viewpoint of intergenerational equity, to require $z^t = z^{t'} < \bar{z}$ for any t, $t' \in \mathbb{T}$. Thus, **Proposition 1** implies that such a requirement of intergenerational equity is inconsistent with Pareto efficiency. To be Pareto efficient, the feasible allocation should have temporary selfish allocations, as **Proposition 2** shows for the case of constant H.

17.5 Main theorems

In this section, we argue the fundamental incompatibility between Pareto efficiency and intergenerational equity. First, we focus on the case $(1-\delta)Z^0 \leq \bar{z}$ in the following discussion. This assumption is reasonable, since the case $(1-\delta)Z^0 > \bar{z}$ implies that even the maximal emissions of public bads by all generations decrease the accumulated amount of public bads, which is counterintuitive. Second, as an efficiency requirement, we consider that **NSPE** is more reasonable than **PE**, since **PE** permits selfish allocations. In fact, it is not so desirable even if **PE** and an intergenerational equity axiom are compatible only at selfish allocations, because such allocations do not resolve the issue of negative externality.

Let us examine the compatibility between Pareto efficiency and intergenerational equity in terms of no-envy, given the above two reasonable restrictions. We arrive at the following fundamental impossibility theorem:

Theorem 1: *Suppose* $(1-\delta)Z^0 \leq \overline{z}$*. Then, there is no allocation rule* φ *which satisfies* ***NSPE*** *and* ***NE****.*

Proof: Let us consider an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ such that for any $t, t' \in \mathbb{T}$, $R_t = R_{t'}$. Moreover, in this economy $\mathbf{R}^{\mathbb{T}}$, we will suppose that every generation's preference R_t is not so much sensitive to the change of accumulated public bads.

Case 1: Let us take any Pareto efficient allocation $\mathbf{a} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ in which there exists a generation $\overline{t} \in \mathbb{T}$ such that $(z^{\overline{t}}, w^{\overline{t}}) = (\overline{z}, 0)$ and $(z^{\overline{t}+1}, w^{\overline{t}+1}) \neq (\overline{z}, 0)$ for the generation $\overline{t}+1 \in \mathbb{T}$. The existence of the generation $\overline{t}$ is guaranteed by **Lemma 5** and **Proposition 1**. Consider $Z^{\overline{t}-1}$ and $Z^{\overline{t}}$ in this allocation. Note that $Z^{\overline{t}-1}$ is consumed by the generation $\overline{t}$ as the third component of the vector $x^{\overline{t}}$, while $Z^{\overline{t}}$ is consumed by the generation $\overline{t}+1$ as the third component of the vector $x^{\overline{t}+1}$. Consider

$$Z^{\overline{t}} - Z^{\overline{t}-1} = -(1-\delta) \cdot Z^{\overline{t}-1} + \overline{z}.$$

We will show $Z^{\overline{t}} - Z^{\overline{t}-1} > 0$. Compare $Z^{\overline{t}-1}$ with $\frac{\overline{z}}{1-\delta}$. Note

$$Z^{\overline{t}-1} = \delta^{\overline{t}-1} Z^0 + \sum_{t'=1}^{\overline{t}-1} \delta^{\overline{t}-1-t'} z^{t'},$$

$$\text{while } \frac{\overline{z}}{1-\delta} = \delta^{\overline{t}-1}\overline{z} + \sum_{t'=1}^{\overline{t}-1} \delta^{\overline{t}-1-t'}\overline{z} + \frac{\delta^{\overline{t}}\overline{z}}{1-\delta}.$$

Also note that

$$\begin{aligned}
\delta^{\overline{t}-1}\overline{z} + \frac{\delta^{\overline{t}}\overline{z}}{1-\delta} - \delta^{\overline{t}-1}Z^0 &= \delta^{\overline{t}-1}\left(\overline{z} + \frac{\delta\overline{z}}{1-\delta} - Z^0\right) \\
&= \frac{\delta^{\overline{t}-1}}{1-\delta}[(1-\delta)\overline{z} + \delta\overline{z} - (1-\delta)Z^0] \\
&= \frac{\delta^{\overline{t}-1}}{1-\delta}[\overline{z} - (1-\delta)Z^0] \geq 0
\end{aligned}$$

by the assumption. Thus, $\frac{\overline{z}}{1-\delta} \geq Z^{\overline{t}-1}$ holds, since

$$\sum_{t'=1}^{\overline{t}-1} \delta^{\overline{t}-1-t'} \overline{z} \geq \sum_{t'=1}^{\overline{t}-1} \delta^{\overline{t}-1-t'} z^{t'}.$$

This implies $Z^{\overline{t}} - Z^{\overline{t}-1} = -(1-\delta) \cdot Z^{\overline{t}-1} + \overline{z} \geq -(1-\delta) \cdot \frac{\overline{z}}{1-\delta} + \overline{z} = 0$. Thus, $Z^{\overline{t}} \geq Z^{\overline{t}-1}$ and $(z^{\overline{t}}, w^{\overline{t}}) = (\overline{z}, 0) \neq (z^{\overline{t}+1}, w^{\overline{t}+1})$ imply that $(x^{\overline{t}}, x^{\overline{t}+1}) \in P(R_{\overline{t}+1})$, since $x^{\overline{t}}$ is temporarily selfish.

Case 2: Let us take any Pareto efficient allocation $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in A$ in which there exists a generation $\overline{t} \in \mathbb{T}$ such that $(z^{\overline{t}}, w^{\overline{t}}) = (\overline{z}, 0)$ and $(z^t, w^t) = (\overline{z}, 0)$ for any $t \in \mathbb{T}$ with $t > \overline{t}$. In this case, compare $x^{\overline{t}}$ with $x^{\overline{t}+1}$. By the same argument on $Z^{\overline{t}} - Z^{\overline{t}-1}$ as Case 1, $Z^{\overline{t}} \geq Z^{\overline{t}-1}$. By the definition of $\mathbf{a}$, $(z^{\overline{t}-1}, w^{\overline{t}-1}) \neq (\overline{z}, 0)$, since $\mathbf{a}$ is non-selfish. In particular, if $z^{\overline{t}-1} < \overline{z}$, then $Z^{\overline{t}} > Z^{\overline{t}-1}$. Thus, $(x^{\overline{t}}, x^{\overline{t}+1}) \in P(R_{\overline{t}+1})$, since $x^{\overline{t}}$ is temporarily selfish. If $z^{\overline{t}-1} = \overline{z}$, then compare $x^{\overline{t}}$ with $x^{\overline{t}-1}$. Since $(z^{\overline{t}-1}, w^{\overline{t}-1}) \neq (\overline{z}, 0)$, $w^{\overline{t}-1} > 0$. Note it may be the case that $Z^{\overline{t}-1} \geq Z^{\overline{t}-2}$. However, in this economy $\mathbf{R}^{\mathbb{T}}$, every generation is not as sensitive to the change of accumulated public bads. Thus, the effect of $w^{\overline{t}-1} > 0$ can cancel out the effect of $Z^{\overline{t}-1} \geq Z^{\overline{t}-2}$, so that $(x^{\overline{t}}, x^{\overline{t}-1}) \in P(R_{\overline{t}-1})$, since $x^{\overline{t}}$ is temporarily selfish.

In summary, if a Pareto efficient allocation is non-selfish, then it does not meet the no-envy condition. Thus, there is no allocation rule which satisfies **NSPE** and **NE.** □

The above theorem implies that any policy for regulating the emissions of public bads and promoting education for human capital is Pareto inefficient whenever it cares about intergenerational equity in terms of no-envy. However, if we give up any such policy, is it possible to implement efficient and equitable allocations in this model? The answer is still negative in general, as the following theorem suggests:

Theorem 2: *Suppose* $(1-\delta)Z^0 \leq \overline{z}$. *Then, there exists an allocation rule* φ *which satisfies selfish-***PE** *and* **NE** *if and only if* $(1-\delta)Z^0 = \overline{z}$.

Proof: Suppose $(1-\delta)Z^0 = \overline{z}$. Then, a historical sequence $(z^{*t}, w^{*t})_{t \in \mathbb{T}}$ with $(z^{*t}, w^{*t}) = (\overline{z}, 0)$ for any $t \in \mathbb{T}$ constitutes a feasible sequence of emitted public bads and investments for human capital, and it holds that $Z^{t-1} = Z^0$ for any $t \in \mathbb{T}$. Given any economy $\mathbf{R}^{\mathbb{T}}$ and this sequence $(z^{*t}, w^{*t})_{t \in \mathbb{T}}$, let us consider a feasible allocation $\mathbf{a}^* = (x^{*t}, z^{*t}, w^{*t})_{t \in \mathbb{T}}$ as follows: for any $t \in \mathbb{T}$, the third component of x^{*t} is given by Z^0; and the first and the second components (l^{*t}, y^{*t}) of x^{*t} are given by: $((l^{*t}, y^{*t}, Z^0), (l^t, y^t, Z^0)) \in R_t$ where (l^t, y^t) satisfies $g(h^0 \cdot l^t, \overline{z}) \geq y^t$. Thus, the allocation $\mathbf{a}^*$ is selfish, so that it is Pareto efficient by **Lemma 1**. Moreover, in this allocation, every generation chooses his selfish action under the same components of the public bads accumulation Z^0 and the human capital accumulation h^0. Thus, $\mathbf{a}^*$ is no-envy.

Consider $(1-\delta)Z^0 < \bar{z}$. Then, in the selfish sequence of emitted pubic bads and investments for human capital $(z^{*t}, w^{*t})_{t\in\mathbb{T}}$, $Z^{t-1} > Z^0$ holds for any $t \in \mathbb{T}\backslash\{1\}$. Consider an economy $\mathbf{R}'^{\mathbb{T}} \in \mathcal{R}^{\infty}$ such that for any $t, t' \in \mathbb{T}$, $R'_t = R'_{t'}$. Consider any selfish allocation $\mathbf{a}' = (x'^t, z^{*t}, w^{*t})_{t\in\mathbb{T}}$ which is consistent with $(z^{*t}, w^{*t})_{t\in\mathbb{T}}$ in this economy $\mathbf{R}'^{\mathbb{T}}$. Then, every generation $t \in \mathbb{T}\backslash\{1\}$ except the generation 1 in this selfish allocation strictly prefers x'^1 to x'^t under the economy $\mathbf{R}'^{\mathbb{T}}$.

Consider $(1-\delta)Z^0 > \bar{z}$. Then, in the selfish sequence of emitted public bads and investments for human capital $(z^{*t}, w^{*t})_{t\in\mathbb{T}}$, $Z^{t-1} < Z^0$ holds for any $t \in \mathbb{T}\backslash\{1\}$. Consider an economy $\mathbf{R}''^{\mathbb{T}} \in \mathcal{R}^{\infty}$ such that for any $t, t' \in \mathbb{T}$, $R''_t = R''_{t'}$. Consider any selfish allocation $\mathbf{a}'' = (x''^t, z^{*t}, w^{*t})_{t\in\mathbb{T}}$ which is consistent with $(z^{*t}, w^{*t})_{t\in\mathbb{T}}$ in this economy $\mathbf{R}''^{\mathbb{T}}$. Then, $(x'^t, x'^1) \in P(R_1)$ holds for any generation $t \in \mathbb{T}\backslash\{1\}$. □

The implication of the above theorem is incompatibility between Pareto efficiency and intergenerational equity in terms of no-envy even over selfish allocations. This is because selfish allocations can be no-envy if and only if $(1-\delta)Z^0 = \bar{z}$, but the occurrence of this equation is almost improbable. In fact, the most probable setting is $(1-\delta)Z^0 < \bar{z}$, which implies the situation that the negative externality to the future generation becomes more serious whenever the current generation emits the maximal amount of public bads.

If the no-envy axiom is replaced by the responsibility and compensation axioms, is it possible to have a better result? Unfortunately, the following theorem gives us a negative answer:

Theorem 3: *Suppose* $(1-\delta)Z^0 \leq \bar{z}$. *Then, there is no allocation rule* φ *which satisfies* ***NSPE*** *and* ***EWUP****.*

Proof: Let us consider an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ such that for any $t, t' \in \mathbb{T}$, $R_t = R_{t'}$. Moreover, in this economy $\mathbf{R}^{\mathbb{T}}$, we suppose the following type of preference:

(i) every generation's preference R_t is not so much sensitive to the change of accumulated public bads;
(ii) every generation's preference R_t meets the *boundary condition* in the sense that for any (l^t, Z^t), $(\tilde{l}^t, \tilde{Z}^t) \in [0, \bar{l}] \times \mathbb{R}_+$, and for any $y^t \in \mathbb{R}_{++}$, $((l^t, y^t, Z^t), (\tilde{l}^t, 0, \tilde{Z}^t)) \in P(R_t)$ holds.

Thus, in this economy $\mathbf{R}^{\mathbb{T}}$, any feasible allocation $\mathbf{a} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ satisfying **EWUP** has the property that for any $t, t' \in \mathbb{T}$, $(x^t, x^{t'}) \in I(R_t)$. In other words, **EWUP** is equivalent to **NE** at $\mathbf{R}^{\mathbb{T}}$. Thus, following the proof of **Theorem 1**, we can see that every non-selfish Pareto efficient allocation at $\mathbf{R}^{\mathbb{T}}$ has a pair of generations $t, t' \in \mathbb{T}$ such that $(x^t, x^{t'}) \in P(R_{t'})$, which indicates the violation of **EWUP**. □

Corollary 1: *Suppose* $(1-\delta)Z^0 \leq \bar{z}$. *Then, there is no allocation rule* φ *which satisfies* ***NSPE, EWUP,*** *and* ***RFG***.

Proof: It is obvious from **Theorem 3**. □

The above impossibility result in **Corollary** 1 comes from the inconsistency of **EWUP** with **NSPE**. Thus, the axioms of responsibility and compensation cannot constitute an efficient allocation rule whenever **EWUP** is taken as the weaker variant of the basic axiom of compensation. By the way, if we take **UNEP** as another weaker axiom of the compensation principle, is it possible to make an efficient allocation rule satisfying the principles of responsibility and compensation? The following theorem still gives us a negative answer whenever **NEEG** is required as the basic axiom of responsibility:

Theorem 4: *Suppose* $(1-\delta)Z^0 < \bar{z}$ *and* H *is constant. Then, there is no allocation rule* φ *which satisfies* ***PE, UNEP,*** *and* ***NEEG***.

Proof: Take any feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t\in\mathbb{T}} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ such that there exists two generations $\bar{t}, \bar{t}' \in \mathbb{T}$ such that $\bar{t} < \bar{t}'$ and $Z^{\bar{t}-1} = Z^{\bar{t}'-1}$. Since we consider the case that H is constant, $h^{\bar{t}} = h^{\bar{t}'} = h^0$. W. l.o.g., let us suppose that $(1-\delta)Z^{\bar{t}-1} < \bar{z}$. Let $z^* \equiv (1-\delta)Z^{\bar{t}-1}$.

Consider the following economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$: for any $t \in \mathbb{T}$ such that $t \geq \bar{t}$, his preference R_t is represented by the following utility function u_t: for any $(\tilde{l}^t, \tilde{y}^t, \tilde{Z}^{t-1}) \in [0, \bar{l}] \times \mathbb{R}_+ \times \mathbb{R}_+$,

$$u_t(\tilde{l}^t, \tilde{y}^t, \tilde{Z}^{t-1}) = \tilde{y}^t - g(h^0 \cdot \tilde{l}^t, z^*) - \tilde{Z}^{t-1}.$$

Suppose **a** has the property of **UNEP** and **NEEG**. Thus, $(x^{\bar{t}}, x^{\bar{t}'}) \in I(R_{\bar{t}}) = I(R_{\bar{t}'})$. Moreover, it follows that the generation $\bar{t}$ emits $z^{\bar{t}} = z^*$ in the allocation **a**, which we will show now. Suppose $z^{\bar{t}} > z^*$. Then, $Z^{\bar{t}} > Z^{\bar{t}-1}$. Then, $z^{\bar{t}+1} > z^{\bar{t}}$ by **UNEP**. Thus, $Z^{\bar{t}+1} > Z^{\bar{t}-1}$. So, $z^{\bar{t}+2} > z^{\bar{t}}$ by **UNEP**, which implies $Z^{\bar{t}+2} > Z^{\bar{t}-1}$. By repeating this process up to $\bar{t}' - 1$, we conclude that $Z^{\bar{t}'-1} > Z^{\bar{t}-1}$, which is a contradiction. By applying the similar argument for the case of $z^{\bar{t}} < z^*$, we can arrive at $Z^{\bar{t}'-1} < Z^{\bar{t}-1}$, which is also a contradiction. Thus, $z^{\bar{t}} = z^*$ holds in the allocation **a**.

Note that to make **a** Pareto efficient at $\mathbf{R}^{\mathbb{T}}$, we have to require that each temporary allocation $\mathbf{a}^t$ is at least $(z^t, 0^t)$-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. Thus, the consumption vector $x^{\bar{t}} = (l^{\bar{t}}, y^{\bar{t}}, Z^{\bar{t}-1})$ in the temporary allocation $\mathbf{a}^{\bar{t}} = (x^{\bar{t}}, z^{\bar{t}}, 0^{\bar{t}})$ should have the property that $y^{\bar{t}} = g(h^0 l^{\bar{t}}, z^*)$. Then, by **NEEG**, we will show that for the generation $\bar{t}' \in \mathbb{T}$, $z^{\bar{t}'} = z^{\bar{t}} = z^*$ and $y^{\bar{t}'} = g(h^0 l^{\bar{t}'}, z^*)$ hold.

First, if $z^{\bar{t}'} = z^{\bar{t}} = z^*$ and $y^{\bar{t}'} = g(h^0 l^{\bar{t}'}, z^*)$ in the temporary allocation $\mathbf{a}^{\bar{t}'} = (x^{\bar{t}'}, z^{\bar{t}'}, 0^{\bar{t}'})$, then $(x^{\bar{t}}, x^{\bar{t}'}) \in I(R_{\bar{t}}) = I(R_{\bar{t}'})$ holds and $\mathbf{a}^{\bar{t}'}$ is $(z^{\bar{t}'}, 0^{\bar{t}'})$-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. Second, if $z^{\bar{t}'} > z^{\bar{t}} = z^*$, then to keep $(x^{\bar{t}}, x^{\bar{t}'}) \in I(R_{\bar{t}}) = I(R_{\bar{t}'})$, it follows that $l^{\bar{t}'} < l^{\bar{t}}$. This is because $y^{\bar{t}'} = g(h^0 l^{\bar{t}'}, z^{\bar{t}'}) > g(h^0 l^{\bar{t}'}, z^*) \geq g(h^0 l^{\bar{t}}, z^*) = y^{\bar{t}}$ whenever $l^{\bar{t}'} \geq l^{\bar{t}}$. Then, however, $x^{\bar{t}'}$ cannot be $(z^{\bar{t}'}, 0^{\bar{t}'})$-constrained temporary Pareto efficient at $\mathbf{R}^{\mathbb{T}}$, since

$$\begin{aligned} u_{\bar{t}'}(\widetilde{l}^{\bar{t}'}, \widetilde{y}^{\bar{t}'}, Z^{\bar{t}'-1}) &= g(h^0 \widetilde{l}^{\bar{t}'}, z^{\bar{t}'}) - g(h^0 \widetilde{l}^{\bar{t}'}, z^*) - Z^{\bar{t}'-1} \\ &> y^{\bar{t}} - g(h^0 l^{\bar{t}}, z^*) - Z^{\bar{t}'-1} = u_{\bar{t}'}(l^{\bar{t}'}, y^{\bar{t}'}, Z^{\bar{t}'-1}) \end{aligned}$$

for any $\widetilde{l}^{\bar{t}'} \geq l^{\bar{t}}$. Third, if $z^{\bar{t}'} < z^{\bar{t}} = z^*$, then $(x^{\bar{t}}, x^{\bar{t}'}) \in P(R_{\bar{t}}) = P(R_{\bar{t}'})$ holds. Thus, **NEEG** implies that $z^{\bar{t}'} = z^{\bar{t}}$.

The above argument implies that $Z^{\bar{t}'-1} = Z^{\bar{t}'}$, so that the property of **NEEG** should be applied for the generations $\bar{t}'$ and $\bar{t}' + 1$. Thus, following the above argument, we arrive at $z^{\bar{t}'+1} = z^{\bar{t}'} < \bar{z}$. In such a way, $z^t = z^{\bar{t}'} < \bar{z}$ holds for any $t \in \mathbb{T}$ with $t > \bar{t}'$, since **a** has the property of **NEEG**. By **Proposition 1**, **a** cannot be Pareto efficient. □

Let us also examine the compatibility between Pareto efficiency and intergenerational equity in terms of egalitarian-equivalence. Unfortunately, the following theorem gives us a negative answer:

Theorem 5: *Suppose* $(1-\delta)Z^0 < \bar{z}$ *and let* $z^* \equiv (1-\delta)Z^0$. *Then, there is no allocation rule* φ *which satisfies* ***PE*** *and* z^*-***EE***.

Proof: Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, let us take any feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t\in\mathbb{T}} = (x^t, z^t, w^t)_{t\in\mathbb{T}} \in A$ which is a z^*-egalitarian equivalent allocation. Thus, for any $t \in \mathbb{T}$ and any $\mathbf{a}^t$, there exists a corresponding ideal allocation $\mathbf{a}_0^t = (x_0^t, z^*, 0^t)$ such that $x_0^t \equiv (l_0^t, g(h^0 l_0^t, z^*), Z^0)$ is the maximizer of R_t under the constraint (Z^0, h^0, z^*). Since $z^* = (1-\delta)Z^0$, the allocation $\mathbf{a}_0 = (\mathbf{a}_0^t)_{t\in\mathbb{T}} = (x_0^t, z^*, 0^t)_{t\in\mathbb{T}}$ becomes a feasible allocation. Since $z^* < \bar{z}$, the feasible allocation $\mathbf{a}_0 \in A$ is $(z^*, 0)$-constrained Pareto efficient at $\mathbf{R}^{\mathbb{T}}$, but not Pareto efficient at $\mathbf{R}^{\mathbb{T}}$ by **Proposition 1**. Thus, there exists an alternative feasible allocation $\mathbf{a}' \in A$ which Pareto-dominates $\mathbf{a}_0$. By the way, since **a** is Pareto indifferent to $\mathbf{a}_0$, **a** is Pareto-dominated by $\mathbf{a}'$. □

Finally, we can also obtain the incompatibility between Pareto efficiency and environmental sustainability. The first incompatibility relevent to **PBM** is given by the following theorem:

Theorem 6: *Suppose* $(1-\delta)Z^0 < \bar{z}$. *Then, there is no allocation rule* φ *which satisfies* ***PE*** *and* ***PBM***.

Proof: Given an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$, let us take any feasible allocation $\mathbf{a} = (\mathbf{a}^t)_{t \in \mathbb{T}} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in A$ which has a monotone non-increasing sequence of public bads consumptions $(Z^{t-1})_{t \in \mathbb{T}}$. Since $(1-\delta)Z^0 < \overline{z}$, and $(Z^{t-1})_{t \in \mathbb{T}}$ is a monotone non-increasing sequence, $z^t < \overline{z}$ holds for any $t \in \mathbb{T}$. Thus, by **Proposition 1**, **a** cannot be Pareto efficient at $\mathbf{R}^{\mathbb{T}}$. □

The second incompatibility is relevent to **J-HD**.

Theorem 7: *Suppose* $(1-\delta)Z^0 < \overline{z}$. *Then, there is no allocation rule* φ *which satisfies* **NSPE** *and* **J-HD**.

Proof: Let us consider an economy $\mathbf{R}^{\mathbb{T}} \in \mathcal{R}^{\infty}$ such that for any $t, t' \in \mathbb{T}$, $R_t = R_{t'}$. Moreover, in this economy $\mathbf{R}^{\mathbb{T}}$, we will suppose that every generation's preference R_t is consistent with J in the sense that for any $x^{t'}, x^{t''} \in X$, $(x, x') \in R_t$ holds if and only if there exists $h \in \mathbb{R}_+$ such that $((x, h), (x', h)) \in J$.

Let us take any Pareto efficient allocation $\mathbf{a} = (x^t, z^t, w^t)_{t \in \mathbb{T}} \in A$ in which there exists a generation $\overline{t} \in \mathbb{T}$ such that $(z^{\overline{t}}, w^{\overline{t}}) = (\overline{z}, 0)$. Then, $h^{\overline{t}} = h^{\overline{t}+1}$. This case corresponds to either **Case 1** or **Case 2** in the proof of **Theorem 1**. Without loss of generality, let us assume that $(x^{\overline{t}}, x^{\overline{t}+1}) \in R_t$ if and only if $((x^{\overline{t}}, h^{\overline{t}}), (x^{\overline{t}+1}, h^{\overline{t}})) \in J$, and $(x^{\overline{t}+1}, x^{\overline{t}}) \in R_t$ if and only if $((x^{\overline{t}+1}, h^{\overline{t}}), (x^{\overline{t}}, h^{\overline{t}})) \in J$ for this $h^{\overline{t}}$. Consider $Z^{\overline{t}-1}$ and $Z^{\overline{t}}$ in this allocation as in the proof of **Theorem 1**. Then, since $(1-\delta)Z^0 < \overline{z}$, $\frac{\overline{z}}{1-\delta} > Z^{\overline{t}-1}$ holds, and so $Z^{\overline{t}} - Z^{\overline{t}-1} = -(1-\delta) \cdot Z^{\overline{t}-1} + \overline{z} > -(1-\delta) \cdot \frac{\overline{z}}{1-\delta} + \overline{z} = 0$. Thus, $Z^{\overline{t}} > Z^{\overline{t}-1}$ and $(z^{\overline{t}}, w^{\overline{t}}) = (\overline{z}, 0)$ imply that $((x^{\overline{t}}, h^{\overline{t}}), (x^{\overline{t}+1}, h^{\overline{t}+1})) \in P(J)$, since $h^{\overline{t}} = h^{\overline{t}+1}$ and $x^{\overline{t}}$ is temporarily selfish for $R_{\overline{t}}$. This implies that **a** violates **J-HD**. □

This theorem implies that non-selfish Pareto efficiency leads to the violation of human development in terms of the objective welfare measure J.

17.6 Concluding remarks

The main theorems put forward in section 17.5 indicate that Pareto efficiency is not so attractive in this context of resource allocations. In this model, the more efficient production of private goods by one generation involves the more emission of the public bad, from which this generation does not suffer. Thus, from the point of this generation's rational choice, they have no motivation to regulate the emission of the public bad. However, from the point of sustainability of human beings as well as the point of intergenerational equity, each generation should implement some policy for regulating the public bad emissions. In contrast, Pareto efficiency requires that there should be generations who never implement any policy for regulating the public bad emissions. Facing these two mutually, incompatible judgements, I believe that the judgement derived from the axioms of sustainability and intergenerational equity should be given priority over the judgement derived from

Pareto efficiency. So, at the expense of Pareto efficiency, we should consider the existence of second-best allocation rules which meet the axioms of sustainability and intergenerational equity as well as the second best efficiency axiom, that would be an open question.

It would be worth commenting on another type of intergenerational equity. Here, we have only discussed the types of welfaristic equity axioms, where the main informational basis for measuring individuals' well-being was individuals' subjective preferences. However, it is possible to discuss intergenerational equity by adopting some *objective measure of well-being*. For instance, we may consider the 'J-Reference Maximin principle' as an intergenerational equity axiom, where J was introduced in axioms of sustainability. Then, it would be interesting to consider the compatibility of the J-Reference Maximin principle with J-Reference Human Development in this context.[7]

Notes

1. Asheim, Buchholz, and Tunggodden (2001) and Asheim and Buchholz (2005) considered a similar type of intergenerational resource allocations to ours, although they discussed the bequest of natural resources, which is characterized as positive externality rather than negative externality. Their economic models are much simpler than ours: they do not have the component of educational investment that our model has.
2. Asheim *et al.* (2001), Asheim and Buchholz (2005), and Roemer (2007) also discuss intergenerational equity and sustainability in production economies with externality, non-overlapping generation, and non-utility discounting. As intergenerational equity criteria, Asheim *et al.* (2001) and Asheim and Buchholz (2005) discuss the Suppes-Sen grading principle and Roemer (2007) takes the 'maximin welfare' principle, while this chapter starts from discussing the no-envy and the egalitarian equivalence principles.
3. In what follows, $\mathbb{R}_+$, $\mathbb{R}^n_+$ and $\mathbb{R}^n_{++}$ denote, respectively, the set of non-negative real numbers, the non-negative orthant and the positive orthant in the Euclidean n-space.
4. Note that the vector inequalities are defined as follows: for any $\mathbf{a}, \mathbf{b} \in \mathbb{R}^q$ with $q > 1$, $\mathbf{a} \geqq \mathbf{b}$ if and only if $a_i \geq b_i$ for all $i = 1, 2, \ldots, q$; $\mathbf{a} > \mathbf{b}$ if and only if $\mathbf{a} \geqq \mathbf{b}$ and $a_i > b_i$ for some $i = 1, 2, \ldots, q$; and $\mathbf{a} \gg \mathbf{b}$ if and only if $\mathbf{a} \geq \mathbf{b}$ and $a_i > b_i$ for all $i = 1, 2, \ldots, q$.
5. Asheim *et al.* (2001), Asheim and Buchholz (2005), and Roemer (2007) adopted this approach for defining sustainability. See also Silvestre (2002; 2007).
6. To define the continuity of the mapping $\triangle \mathbf{z}^{\mathbb{T}}$, we may adopt the *sup metric* as the topology of $\mathbb{R}^{\infty}_+$.
7. A similar type of problem was addressed and solved by Roemer (2007) and Silvestre (2007) in different types of models respectively.

References

Asheim, G. B. and W. Buchholz (2005) 'Can Stock-specific Sustainability Constraints be Justified?', in S. Kant and R. A. Berry (eds), *Sustainability, Economics and*

Natural Resources: Economics of Sustainable Forest Management (New York: Springer), pp. 175–89.

Asheim, G. B., Buchholz, W. and B. Tungodden (2001) 'Justifying Sustainability', *Journal of Environmental Economics and Management*, vol. 41, pp. 252–68.

Dworkin, R., (1981) 'What is Equality? Part 2: Equality of Resources', *Philosophy and Public Affairs*, vol. 10, pp. 283–345.

Fleurbaey, M. (1994) 'On Fair Compensation', *Theory and Decision*, vol. 36, pp. 277–307.

Fleurbaey, M. (1995) 'Three Solutions to the Compensation Problem', *Journal of Economic Theory*, vol. 65, pp. 505–21.

Fleurbaey, M. and F. Maniquet (1996) 'Fair Allocation with Unequal Production Skills: the No-Envy Approach to Compensation', *Mathematical Social Sciences*, vol. 32, pp. 71–93.

Foley, D. (1967) 'Resource Allocation and the Public Sector', *Yale Economic Essays*, vol. **7**, pp. 45–98.

Pazner, E. and D. Schmeidler (1978) 'Egalitarian Equivalent Allocations: A New Concept of Economic Equity', *Quarterly Journal of Economics*, vol. 92, pp. 671–87.

Roemer, J. E. (2007) 'Intergenerational Justice and Sustainability, under the Leximin Ethic', Ch. 12 in this volume.

Sen, A. K. (1980) 'Equality of What?', in S. McMurrin (ed.), *Tanner Lectures on Human Values*, vol. 1 (Cambridge: Cambridge University Press).

Sen, A. K. (1985) *Commodities and Capabilities* (Amsterdam: North-Holland).

Silvestre, J. (2002) 'Progress and Conservation under Rawls's Maximin Principle', *Social Choice and Welfare*, vol. 19, pp. 1–27.

Silvestre, J. (2005) 'Intergenerational Equity and Human Development', Ch. 14 in this volume.

Suzumura, K. and K. Tadenuma (2007) 'Normative Approaches to the Problem of Global Warming: Responsibility and Compensation', Ch. 16 in this volume.

Index