572 LECTURE NOTES IN ECONOMICS AND MATHEMATICAL SYSTEMS

Matthias F. Jäkel

Pensionomics

On the Role of PAYGO in Pension Portfolios

Springer

Lecture Notes in Economics and Mathematical Systems 572

Matthias F. Jäkel

Pensionomics

On the Role of PAYGO in Pension Portfolios

With 17 Figures
and 10 Tables

Springer

Matthias F. Jäkel

WHU, Otto Beisheim School of Management
Dresdner Bank Chair of Finance
Burgplatz 2
56179 Vallendar
Germany
E-mail: mjaekel@whu.edu

Cataloging-in-Publication Data
Library of Congress Control Number: 2006923635

ISBN-10 3-540-32597-2 Springer Berlin Heidelberg New York
ISBN-13 978-3-540-32597-0 Springer Berlin Heidelberg New York

Springer is a part of Springer Science+Business Media
springeronline.com

Printed in Germany

Cover design: Erich Kirchner
Production: Helmut Petri
Printing: Strauss Offsetdruck

SPIN 11678649 Printed on acid-free paper – 88/3153 – 5 4 3 2 1 0

To my parents

Foreword

This thesis is the result of my doctoral studies at the Dresdner Bank Chair of Finance at the WHU – Otto Beisheim School of Management. The WHU, a young, private institution and trailblazer in the German educational landscape, played a central role in my academic, professional and personal development. Having completed my undergraduate and graduate studies in this friendly and informal but at the same time encouraging and challenging atmosphere, I was more than happy to continue my studies at the doctoral level in WHU's unique setting.

Yet such an academic enterprise could not take place without the generosity of many selfless people who give hours of their time to listen to, engage with, and improve upon the works of others. Therefore, I want to especially thank my supervisor and mentor Professor Dr. Markus Rudolf for his guidance and support throughout my time at the Dresdener Bank Chair of Finance. Early in my final study year he drew my attention to financial research, and my interest into the subject of demography and finance can also be traced back to my master thesis produced under his supervision. During my time as a doctoral candidate, he provided very helpful instruction and useful advice as well as encouragement and constructive criticism when needed. At the same time he understood to generate a friendly and harmonious environment with sufficient academic freedom for his team of young researchers.

I thank Professor Dr. Michael Frenkel for acting as co-supervisor for my dissertation. His challenging but always interesting and insightful courses throughout my studies at the WHU were the cornerstone in developing my own critical economic thinking. He provided an invaluable second perspective on the subject and was a great help in challenging and validating the path I had chosen.

Furthermore, I would like to deeply thank my colleagues at the Dresdner Bank Chair of Finance – Miriam Begtasevic, Marianne Diel, Michael Adams, Volker Anger, Dr. Frank Guse, Dr. Karl Keiber, Dr. Matthias Muck, Valentin Ulrici, and Marco Vietor – for making it an extremely enjoyable and productive environment. I had a wonderful experience there and will always have very pleasant memories of my time in Vallendar.

This work and my general knowledge of financial economics have benefited greatly from my time as visiting student in the PhD program at the University of Chicago Graduate School of Business. I am thankful to Professor George Constantinides and to Professor John Cochrane for offering me their challenging criticism on my ideas. I gained most of my insights into demography in Professor Robert Fogel's course "Population and the Economy". Malaina Brown was very supportive in organizing my visit.
I am indebted to Jans Meckel – my old partner in academic crime – as well as Daniel Pindur for proofreading and providing valuable comments on the pre-final versions. Furthermore, I would like to express my gratitude to the staff of Springer-Verlag, in particular Katharina Wetzel-Vandai and Christiane Beisel, for their support and help in the realization of the publication.
I cannot overestimate the impact of Dr. Astrid Eisenberg on this work. She has not only enriched my start at the Dresdner Bank Chair of Finance but also helped me through all the ups and downs of my doctoral studies. With her ample academic knowledge, supporting constructiveness, and seemingly endless patience she played an important role in substantially improving the content and style of this thesis.
Last but not least, my greatest appreciation goes to my parents Brigitte and Peter Jäkel for their constant encouragement and total support throughout my academic education and all my life. I dedicate this book to them.

Frankfurt, January 2006 *Matthias Jäkel*

Contents

Part V Lists and References

Part I

The Problem

1

Global Aging and Pensionomics

Alice in Fiscal Land

One day a girl named Alice fell down a very deep hole near a tree in the park. After lots of adventures and bumping into strange-looking accountants, actuaries, and tax attorneys, Alice finds herself face to face with Tweedledum and Tweedledee, two citizens of different countries in Fiscal Land. The two are having a heated debate about which country has the better fiscal policy:

Tweedledum: We tax the young $1,000 each to help the old. But we're debt free.
Tweedledee: Well, we also help the old, but we do it by borrowing $1,000 each from the young. And we tax the old $100 each to pay the interest on the debt.
Tweedledum: Well, you have debt. That's bad.
Tweedledee: Well, you tax young people. That's worse.
Tweedledum: Debt erodes one's moral fiber.
Tweedledee: Taxing the young to support the old is exploitive.

Back and forth, back and forth, the two argue, hour after hour, until Alice, who has a Ph.D. in linguistics, screams: "Your countries have exactly the same fiscal policy!" "No way," they both reply. "His has debt, and mine doesn't," says Tweedledum. "And his taxes the young and mine taxes the old," yells Tweedledee. "Of course you're right, but actually you're wrong." Alice says as she shows them table [below]. "Look," Alice says, "in both countries, the young hand over $1,000 each to the government, and the old receive, on balance $1,000 from the government. In one country, the government takes $1,000 from the young and calls it a tax. In the other, it calls it borrowing. In one country, the government hands $1,000 to the old and calls it a transfer payment. In the other , government hands $1,000 to the old and calls it a return plus interest of $1,100 less a tax of $100. So you have the same policies; you're just calling them different things."

The two Tweedles stare dumbfounded at Alice and then exclaim: "Oh, we get it. You're an economist. That's why we don't understand a word you're saying." "Trust me," says Alice. "I'm nothing of the kind. This is just a matter of logic. In your country, Tweedledum, young people get back in old age what they pay in taxes when young, but they don't earn any interest. And in your country, Tweedledee, young people get back in old age with interest what they paid in principal when young, but then the interest is taken away from them." "You really are an economist," shout the Tweedles. "You interrupt a critically important policy debate with some cockamamie theory that has no connection to reality. How about getting lost?"
At this, Alice breaks down in tears and runs headlong into the Queen of Hearts, who promptly arrests her for promoting tax evasion.

Kotlikoff and Burns [2004]: "The Coming Generational Storm", p. 74-76

Deficit Delusion in Fiscal Land

Transaction	*Tweedledum's Country*	*Tweedledee's Country*
Net payment when young	*$1,000*	*$1,000*
Net receipt when old	*$1,000*	*$1,000*
Description of net payment when young	*A $1,000 tax payment*	*Purchase of a $1,000 government bond*
Description of net receipt when old	*A transfer payment*	*Payment of $1,100 in principal and interest less a $100 tax*

Pensionomics

In the course of the global demographic change, population aging will be one of the most important and radical changes in human history. While this is not a new insight most of the public debate and scientific research focussed on the direct socio-political consequences only, especially on the sustainability of fiscal regimes and pay-as-you-go pension systems. Yet, the demographic change will fundamentally alter the understanding of the economy as it affects all of its central markets – those for human labor, for goods and services as well as for financial capital. Facing this, the question of how to effectively and efficiently deliver retirement income will be one of the most eminent problems, not only for the Tweedles. For answering it an in-depth combination of pension finance and economics is required: a step into Pensionomics.

1.1 Global Aging

Century-long population growth has given rise to the modern societies and science has helped to generate and to understand the standard of living now enjoyed. Economics has emerged as one of the sciences focusing on managing scarcity – of consumable goods, physical capital or real property. Now an unprecedented change is taking place and will ultimately alter the focus onto scarcity of *human beings*. A short retrospection of the past demographic trend, its impact on economic development and the scientific understanding of it builds some understanding for the upcoming seismic shift in its full breadth and depth.

History of Population Trends

Assessing the consequences of population change on the pace and process of economic growth is one of the oldest topics in the economic literature. The debate began with Malthus [1798] and his proposition that population would grow at a geometric rate while food supply increases at an arithmetic one, which eventually overcomes the world's ability to feed itself and results in widespread starvation. His view of economics as a dismal science can only be understood based on the historic development of global population.

During their first 100,000 years homo sapiens increased their population at an extremely slow rate. Only the discovery of agriculture at about 9000 B.C. made possible a higher population growth, allowed the creation of cities and eventually led to the first great civilizations in Mesopotamia, India, China and ancient Latin America. The moderate increase continued until around A.D. 600 when a millennium of population cycles with periodic pandemics of diseases like the black plague restrained further expansion. Then, at about 1700, the Agricultural Revolution with its enormous advances in food production permitted the population virtually to explode.[1] Global population started a vast increase from about 500 million in 1700 to more than 6 billion in 2000 and experienced dramatic social changes like the Industrial Revolution and the French Revolution. The start of modern economics as a discipline of its own by Smith's [1776] discussion of political economy must be seen in this context.

Mankind escaped the Malthusian doom, because Malthus failed to predict these agricultural and economic advances. His pessimist view was based on the assumption that food supplies were directly proportionate to the amount of land available. Though this had been historically true, the relationship changed exactly when Malthus extrapolated the empirical evidence to make his predictions: the Agricultural Revolution had just enabled English farmers to produce more food from the same area of land. So Malthus observed the resulting faster population growth but contrasted it with the ancient development in food production. In the long run, he was wrong, too, because the population change he considered geometrical turned out to be sigmoidal. Yet, the leveling-out of this trend takes more than two centuries and is not only driven by the birth rate.

The resulting change of population has been coined *Demographic Transition* based on an interpretation of Thompson [1929].[2] While the more developed countries have been experiencing this transition over the past 200 or so years, less developed countries only began it in the 20th century, yet at a considerably faster rate of change. The theoretic model for the Demographic Transition

[1] This mostly follows Fogel and Costa [1997], whose Table 1 highlights the dramatic growth in population started in 1700. Currais [2000] and Weir [1987] provide deeper understanding on the history of population economics.

[2] The theory of demographic transition has been substantially formed by Notestein [1945] and enlarged by Blacker [1947].

is based on changes in birth and death rates which typically can be categorized in four stages: The first characterizes the pre-modern times before the Agricultural Revolution with balancing birth and death rates at very high levels. This was the Malthusian stalemate with hardly any change in population. Stage two sees a skyrocketing rise in population caused by a decline in the mortality, while fertility remains still high out of tradition and practice. The decline in the death rate, particularly in childhood, is due to the improvements in food supply and to the significant progress in public health – less based on medical breakthroughs than based on advances in water supply, sewage, food handling, and general personal hygiene. The result is another key characteristic of the second stage: the age structure of the population becomes increasingly youthful creating the now famous age pyramid. However, industrialization and urbanization altered the traditional values placed upon fertility and changed the value of children in society. This gives the third stage of the Demographic Transition that moves the population towards stability through a decline in the birth rate. The ideas of the New Home Economics replacing the Malthusian perspective help to understand this change in fertility: starting with Becker [1960] economists view parenthood as driven by the consumption-like pleasure of having children and the provision of future material benefit to the family unit. The parental demand for children depends hence on family income, how parents value their time and on quality-like characteristics of children.[3] Yet, even at the end of the third stage, when fertility has dropped to the level of replacement, population growth continues for a while due to population momentum, i.e. a relatively high concentration of people in the childbearing years. Finally, the Demographic Transition ends with stage four, which is again characterized by stability. However, frequently the fertility rate falls well below the replacement level and population decline sets in rapidly.

While many less developed countries are still in the second stage, most industrialized countries would have actually finished the transition in the late 20th century. However, following World War II Europe, North America and to a lesser extent also Japan experienced a massive increase in fertility. This *Baby Boom* lasted until the mid 1960s when birth rates dropped back. While conventional wisdom links the Baby Boom to the general optimism following the soldiers' return from World War II and the prevailing economic boom, a new explanation of Greenwood et al. [2005] accredits it to the invention of labor-saving household products and management techniques. Since the end

[3]Other major contributions are Easterlin [1968], Becker and Lewi [1973], and Becker [1991]. Willis [1987] provides a summary on New Home Economics. The formation of "quality" children, i.e. qualified workers, gave rise to the emphasis on human capital *formation*; see Schultz [1961, 1963, 1974], Becker [1962, 1993], or Becker and Tomes [1986]. Barro and Becker [1989], Becker et al. [1990], Rosenzweig [1990] and others linked this to economic growth. As will be argued later the view of investment in human capital in the sense of *forming* qualified workers is *not* addressed here.

of the Baby Boom, fertility rates have not only dropped to pre-boom levels, but also continued a decline to sub-replacement levels.
In the early second half of the 20th century, a new concern emerged as developing countries experienced exceptionally high population growth rates based on boom-like fertility and declining mortality. As it was unclear whether the predictable fertility decline a la Demographic Transition would be fast enough to avoid potentially deteriorating effects on economic progress and the environment, the pessimistic Malthusianism reappeared: the famous study for the Club of Rome by Meadows et al. [1972, 1974] predicted uncontrollable mass starvation again based on extrapolation of observed trends – just like Malthus [1798]. Though Asian countries rather succeeded with the implementation of the recommended birth controls, concerns on the Asian population growth continued.[4] Yet, again agricultural and technological advances helped to avoid the doom.

Future Population Trend

Given this understanding of the long-term trend, the current outlook on the global population is hardly surprising. It is deeply driven by changes due to the Demographic Transposition. As illustrated in Table. 1.1 the world is expected to have 6.5 billion inhabitants by mid 2005. Despite the declining fertility levels projected over 2005-50, most forecasters see a peak of world population at about 9.1 billion.[5]
The reason for this leveling-out is the sharp fall in fertility in course of the Demographic Transition. As seen in Table 1.1, women in *all* regions have now less children than in the 1970s and this decline is expected to continue. For China its one-child policy has accelerated the fertility decline so much that its fertility is estimated at 1.7 children per woman during 2000-05 – far below the 2.1 replacement level. The table gives also data for the other two main drivers of population: mortality and immigration. Global life expectancy at birth – risen from 46 years in 1950-55 to 65 years in 2000-05 – is expected to keep on rising and to reach 75 years in 2045-50. Since the movement of people across international boundaries is very often a response to changing socioeconomic, political and environmental forces, international migration is subject to a great volatility and the component of population change most difficult to define, measure and estimate. Still, the United Nations project a net number of 98 million international migrants to more developed regions for the period 2005-50 – an average of 2.2 million annually.[6]

---

[4] See for instance Brown [1992].

[5] Figures are based on United Nations [2004*a*, p. 3] and United Nations [2005*a*, p. 6]. Other international sources on population forecasts are United Nations [1999], Donaldson [1999], Haupt and Kane [2004] or U.S. Census Bureau [2004]. Pötzsch and Sommer [2003] focuses on Germany and Day [1996] on the United States.

[6] For details on all aspects of population projections see United Nations [2005*a*,*b*,*c*,*d*], where all the demographic figures are taken from.

Table 1.1. World population prospects

	Population (millions)			Life expectancy at birth	
	1975	2005	2050	2000 – 05	2045 – 50
World	4,074	6,465	9,076	64.7	74.7
More developed regions	1,047	1,211	1,236	74.6	81.7
Less developed regions	3,027	5,253	7,840	62.8	73.6
Africa	416	906	1,937	48.8	65.3
Asia	2,395	3,905	5,217	67.0	77.1
Europe	676	728	653	73.3	80.4
Latin America	322	561	783	71.0	79.4
Northern America	243	331	438	77.4	82.6
Oceania	21	33	48	74.0	81.2

	Total fertility (children per woman)			Net immigration (thousands p.a.)	
	1970 – 75	2000 – 05	2045 – 50	2000 – 10	2040 – 50
World	4.49	2.65	2.05	n.m.	n.m.
More developed regions	2.12	1.56	1.84	2,462	2,158
Less developed regions	5.44	2.90	2.07	−2,462	−2,158
Africa	6.72	4.97	2.52	−410	−322
Asia	5.08	2.47	1.91	−1,244	−1,204
Europe	2.16	1.40	1.83	937	699
Latin America	5.05	2.55	1.86	−740	−567
Northern America	2.01	1.99	1.85	1,360	1,300
Oceania	3.23	2.32	1.92	98	94

Remarks: The table gives historic data and projections on the global population and its driving demographic variables. Latin America includes the Caribbean; see United Nations [2005*a*, pp. 153-155] for a description of the areas and regions. Medium-variant is used for 2050-projections.
Source: United Nations [2005*a*, Tables I-1, II-1, III-1, IV-1].

For the developed world, this level of net migration will mostly offset the expected excess of deaths over births during 2005-50, which itself amounts to a loss of 73 million people. On the other hand, the 98 million emigrants represent hardly less than 4% of expected population growth for the developing world. Already during 2000-05, net migration in almost thirty countries either prevented population decline or doubled at least the natural contribution to population growth. Among these countries are Austria, Canada, Denmark, Germany, Italy, Portugal, Spain, Sweden, and the United Kingdom. The United States of America, Germany and Canada are projected to be the three major net receivers of international migrants with annually averages of 1.1 million, 204,000 and 201,000, respectively. Still, for more then 50 countries, including Germany, Italy, Japan, the Baltic States and most of the members of the Commonwealth of Independent States, national population is expected to be lower in 2050 than in 2005.

The direct consequence of these trends is population *aging*, whereby the share of older persons in a population increases relative to that of younger persons. But global aging is not only due to the decline in fertility but also driven by the increases in life expectancy. Globally, the number of persons aged 65 years or over is expected almost to triple, increasing from 475 million in 2005 to nearly 1.5 billion by 2050. Even more pronounced is the expected increase in the number of the oldest-old, i.e. persons aged 80 years or over: from 86 million in 2005 to 394 million in 2050.

Figure 1.1 illustrates the aging effect for the six major world regions. Asia has, of course, the most old people in absolute numbers. The *old-age dependency ratio* is the most widely used measure to correct for this size effect describing a population's demographic composition with respect to age. It is the number of persons aged 65 or older per one hundred persons of age 15 to 64.[7]

Although current regional differences in the old-age dependency ratio are expected to persist well into the next 50 years, all six major areas will experience remarkable growth in this ratio over the next half-century: the world's highest old-age dependency ratio of Europe will almost triple in the next 50 years. Based on the projections of the United Nations [2004*b*, 2005*a*], the ratios of Asia and Latin America will more than triple, those of Northern America and Oceania almost double and even Africa's ratio increases by two thirds. While the world's aggregate dependency had increased only from 9.6 in 1950 to 11.4 in 2005, it is expected to reach 25.4 in 2050. In developed countries the elderly population has already surpassed the number of persons aged 14 or less and

[7] Other definitions of the ratio, sometimes also referred to as old-age ratio, include the population of ages between 15 and 59 only. In this sense the figures here are rather conservative. The old-age dependency ratio and the youth dependency ratio – persons under 15 per persons aged 15-64 – sum to the total dependency ratio. These indicators give insight into the amount of people of non-working age compared to the amount of people of working-age. Another measure for the pure aging effect is the age-index as the number of persons of 60 years or older per hundred persons under age 15. For more definitions of indicators see United Nations [2005*a*, p. 41-42].

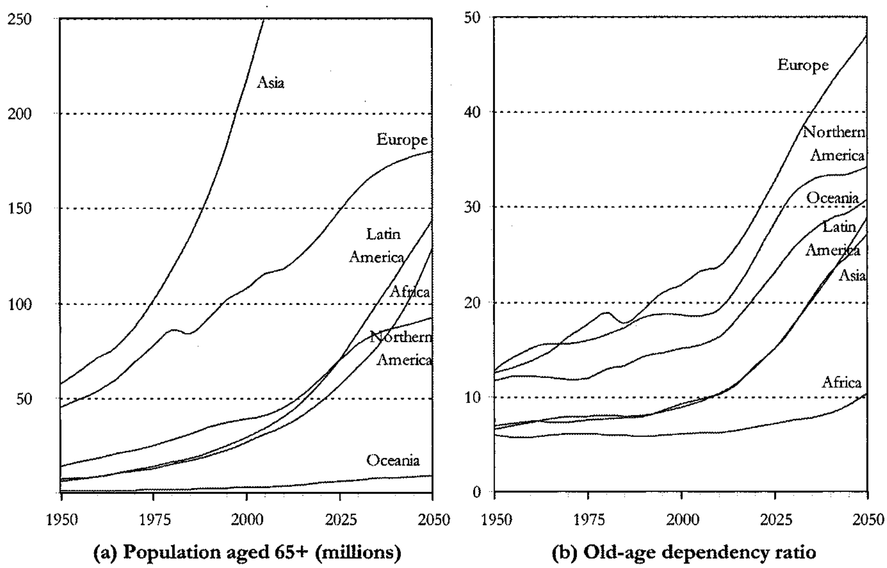

Fig. 1.1. Global aging

Remarks: Graph (a) shows the number of people aged 65 and older for the six major world areas. Asia will exceed 900 millions in 2050. Graph (b) depicts the corresponding old-age dependency ratio, that is the number of persons aged 65+ per one hundred persons of age 15 to 64. Latin America includes the Caribbean.
Data Source: United Nations [2004*b*, 2005*a*], Medium-variant of population database.

by 2050 there will be even 2 elderly persons for every child. While today only 11 developed countries have a median age above 40 years, by 2050, there will be 90 countries in that group. For the world, the median age is projected to increase by about 10 years over the next half century, having only risen by about 3 years in the past 50 years. Population aging – in developed countries already a pervasive reality – is also inevitable in the developing world and will occur there considerably faster. Aging is thus not limited to any region or country, but a truly *global phenomenon.*[8]
The Demographic Transition as demography's mainstream theory perfectly explains the observed changes and is the theoretical foundation for these projections: due to the social and economic development, mortality declines globally, resulting in a higher life expectancy as seen in Fig. 1.1. Fertility

[8]Further indicators of past and future aging can be found in United Nations [2001] or Kinsella and Velkoff [2001], but also general population forecast like United Nations [2005*a*] emphasize this issue. Birg [2003] provides an overview over the subject focussing on Europe and Germany.

follows this trend, yet with a time lag, inducing a substantial population growth. Whereas developed countries took centuries for this process, some rapidly developing countries like the Economic Tigers are now transforming in mere decades. Hence, the more developed regions – in particular Europe – have been leading population aging, but all regions will experience the same changes. While industrialized countries, above all the United States, managed to compensate the falling birth rates with immigration from less developed regions, the actual aging will become more noticeable in the future: it would not only require ever greater numbers of immigrants, but the situation will also be aggravated as the baby-boomers aging. While so far this demographic extra-effect helped to reduce the old-age dependency ratio based on disproportionately many persons of age 15 to 64, the result of the Baby Boom will drastically deteriorate the old-age dependency ratio in the period 2010-25, when the large cohorts not only leave the working age but also enter the group of 65+. This explains the clear steepening of Europe's line for the old-age dependency ratio in Fig. 1.1.

In accordance with the Demographic Transition proposition, the phase of highest population growth is over. The increase in world population, which slowly started in the 17^{th} and 18^{th} centuries and rose dramatically in course of the 20^{th} century, reached a peak at 2% per year in 1965-70. While the studies of Meadows et al. [1972, 1974] extrapolated this trend, the actual growth rate of the world population has declined since then, reaching 1.2% per annum in 2000-05, and is expected to drop further to an annually 0.4% by 2045-50. The sigmoidal population trend is leveling-out after 350 years.

Concerning future fertility and mortality, two more concepts must be mentioned, which currently influence demographic research. Among population scholars the concept of a "Second Demographic Transition" has become an important additional explanation concerning the demographic change in European societies. This theory stresses the importance of ideational changes for altering demographic behaviors such as single living, pre- and post-marital cohabitation, delayed fertility or high prevalence of non-marital fertility.[9] Concerning mortality patterns, the process of the "Technophysio Evolution" has been introduced by Fogel [1994] as a synergism between technological and physiological improvements leading to greatly improved robustness and capacity of the human physiology. Though this process started with the Agricultural Revolution and has since then increased body size by over 50%, it is still going on and hence relevant for forecasting longevity and morbidity.[10] Similar ideational changes on a global level or continued physiological advance will only aggravate the phenomenon of global aging.

[9] The concept of the Second Demographic Transition has been introduced by Lesthaeghe and van de Kaa [1986]. See also Lesthaeghe and Neels [2002] and Lesthaeghe and Surkyn [2004].

[10] See also Fogel [1997] and Fogel and Costa [1997]. Fogel [2003] relates the changes of the Technophysio Evolution to health care costs and intergenerational conflicts. A summary of the longevity issue is found in Vaupel [1998].

To summarize, within the next 50 years a three centuries old trend of population growth and average rejuvenating will revert globally. For the industrialized countries this global aging is intensified by the long-term consequences of the post-war Baby Boom.

1.2 Pension Problem

Unlike other predictions about the future, global aging is not a mere hypothesis. As its approximate timing and magnitude are already locked in, it is like a revolution sure to happen. Yet, it is important to see population aging as a *human success story* in the first place: the triumph of public health, medical advancements, and economic development over diseases and injuries that limited human life expectancy for millennia. However, one must also acknowledge that the phenomenon of aging will bring dramatic consequences: global aging will restructure economies, reshape families, redefine politics and even rearrange the geopolitical order of the next century. The focus of this contribution will be only on the economic development and on the challenges which aging poses on the way of providing and sustaining retirement income for the elderly. Though this particular question is certainly not the only pressure global aging poses, is is definitely one of the most urgent ones.

Economic Challenges of Aging

It is widely agreed that global aging has a deep impact on macroeconomic development. The main channels, through which aging affects the economy, are changes in the labor force, the effect on public finances from increased pressure on public health and pension expenditure, and influences on private saving behavior as a result of life-cycle effects. All of these must – at least – be considered in a thorough analysis. Furthermore, as the overall population declines in the coming decades, so will the number of consumers and producers. There will be fewer persons in the household-forming years leading to overcapacity and falling returns on investment in such key sectors as construction, real estate and durable goods. Many industries will lose the possibility to reach economies of scale.[11]
While for many decades demographics and labor force participation trends have provided a favorable economic environment in many developed countries, global aging will start an era of tight labor markets. Though it is hardly

[11]Detailed assessments of the macroeconomic and financial implications of aging populations can be found in Group of Ten [1998], McMorrow and Roeger [1999], OECD [1998] or Turner et al. [1998]; OECD [2000]. The Center for Strategic and International Studies has established an age-vulnerability index for various economies; see Jackson and Howe [2003] and Jackson [2003] for Germany. England [2002*a*,*b*, 2005] provides good summaries on the impact of aging on macroeconomics, financial markets and on the special role of China. China's pronounced aging problem is also the focus of Cheng [2003].

imaginable given present problems on the labor market, the workforce in much of continental Europe and Japan is projected to be contracting by roughly 1% annually by 2010. In the past, the fraction of the population at work had *constantly* been growing, even though generous retirements rules allowed for a dramatic fall in the number of individual years spent in work. Unless the future decline in the workforce is offset by increases in labor productivity and the effective supply and utilization of it, the growth of material living standards will fall. Of course, this concern is based on the assumption that the elderly will hardly contribute to the economy in the future. Yet, governments may respond with policies designed to expand the labor supply by encouraging more citizens to work and by encouraging those who are already employed to work more and longer in their lifetime. However, in most countries such policies, which could soften the economic consequences of the increasing dependency-ratio, have to overcome deep-rooted social expectations — about early retirement, shorter work weeks or the role of women. Moreover, these policies are unlikely to be effective unless governments reform public retirement systems first, whose rising costs present a big disincentive for work. Only sluggishly are governments beginning to accept that *substantial* reform is necessary to ensure the viability of the public systems for social security.
When German Chancellor Otto von Bismarck established the world's first national old age provision plan in 1889, he designed it as contributory "social insurance": by relating benefits to contributions made, this pension plan had elements of an insurance for the participating workers and was not just a pure redistribution scheme. While the basic principle has remained essentially the same, economic circumstances have changed dramatically: in the beginnings of public pension schemes the state promised a modest benefit to a tiny minority of workers lucky enough to survive to the age of 70. Today, the vast majority of Germans can look forward to a retirement lasting a third or more of their adult lives. Worldwide numerous countries have copied the German system. Above all, after the World War II public pensions have become the financial lifeline of the elderly in many societies. Old-age pension schemes are now social institutions in many – if not most – countries throughout the world. The goal of most public old-age pension schemes is to provide all qualifying individuals with an income stream during their later years. In most developed countries more than 90% of the labor force are now covered by mandatory old-age pension plans. The arrangements are commonly referred to as "pay-as-you-go" or PAYGO systems, insofar as current revenues – contributions from current workers – are used to finance the pension payments of people who have already retired from the labor force. Like the German archetype, most PAYGO systems were initially based on a small number of pensioners relative to a large number of contributing workers and could hence promise generous benefits. As systems matured and the Demographic Transition became ever more noticeable, ratios of pensioners to contributors grew. In some countries the PAYGO systems proofed to be unsustainable, particularly during periods of economic stagnation. While part of the measures taken by politics –

or at least considered – include increases in contribution rates, restructuring or reducing of benefits, and raising the standard age of retirement, a major result of the problems was the development of private pension systems to complement the public PAYGO ones.[12]

Crumbling Pillars of Old Age

In particular the landmark report of the World Bank [1994] advocated such a mix of systems. Therein the World Bank identified three main roles of pension systems: poverty relief, consumption smoothing, and redistribution from lifetime rich to those at risk of old-age poverty. The bank concluded that the challenges of global aging could be met best by combining the above mentioned public pay-as-you-go system with a privately managed mandatory saving accounts and voluntary retirement savings. This was the origin of the "multi-pillar" pension model. After many countries in Latin America, Europe and central Asia adopted these proposals, the bank supplements the initial three systems in its new report by Holzmann and Hinz [2005] by two more pillars: a tax-financed safety net and non-financial means such as family support or access to health care. Furthermore, it establishes the goals that pension schemes should achieve: they are supposed to provide adequate, affordable, sustainable, and robust benefits. Adequacy refers to both the absolute level as well as the relative lifetime level of retirement income that the pension system will provide. In order to prevent negative consequences for individual economic lifetime opportunities and general economic growth, pension schemes must obey individual and societal financing limits and be affordable. Sustainability refers to the current and future financial soundness of the schemes. That is, the pension system should be structured so that the financial situation does not require unexpected contribution hikes, unannounced future benefit cuts, or major and unforeseen transfers from the general fiscal budget. Finally, the system is supposed to withstand major shocks and to remain viable under unforeseen circumstances, i.e. it must be robust.
With this combination of various pillars, the World Bank follows *explicitly* the diversification idea from finance. The multi-pillar approach can be viewed as a portfolio-like combination of pension systems to address and manage the risks of aging. This view of pension mechanism as risk management devices implies that their design must be based on an assessment of their capacity to manage the relevant risks on an individual and a collective level. Each pillar is characterized by consumption-allocating elements that can either be viewed

[12]See ISSA [2003*a*,*b*, 2004*a*,*b*] for a detailed description of social security systems in the world. Kinsella and Phillips [2005] gives a detailed breakdown of the challenges of aging in pensions, health care, disability and well-being. Given its importance the social security issue is in the focus of most publications on aging, like Chand and Jaeger [1996], Lee and Skinner [1999], Mackenzie et al. [2001], Holzmann and Stiglitz [2001], European Commission [2001], Kinsella and Velkoff [2001], Jackson [2002], Dunaway and N'Diaye [2004], or Holzmann [2004].

as assets or as liabilities and affect individuals and society – like those of Fiscal Land – in this matter. The advantage of combining them is that the varying characteristics allow to achieve the desired individual and societal benefits, while minimizing the relevant risks. In other words, each individual pillar on its own has its drawbacks and problems.[13]

While the World Bank's pillars may refer to *five* different organizational structures for realizing pension systems institutionally, there are only *two* underlying principal methods of financing them: a *funded* pension systems resembles an investment fund in which the contributions allow to acquire assets used later to pay for the benefits. The described pay-as-you-go mechanisms, on the other hand, is an *unfunded* system, where no assets are set aside, but benefits are paid for by the pension sponsor as and when they are paid. Typically, PAYGO pension arrangements are financed directly from current working generation by wage-based contributions. While the problem with this method is evident given the expected rise in old-age dependency ratios, funded pension pillars face also the deteriorating consequences of global aging.

The so-called "asset market meltdown" as it was coined by Porterba [1998, 2001] is the most prominent hypothesis on demography's impact on asset returns. According to this hypothesis, the clustered retirement and following dis-saving of comparatively large baby-boomer cohorts may cause a stock market crash after this generation's synchronous saving behavior has driven up asset markets for 40 years. This argument extends the view of the much-publicized study of Mankiw and Weil [1989] on demography-driven housing demand. They argue that the age-specific demand will have negative implications for future house prices due to the aging of the population. Other alerters of the meltdown include Siegel [1998, "Sell? Sell to whom?", p. 41] and Brooks [2000], who warns that historically high returns on the funded pension pillars might have been exceptions and will not be observed in the future. The significance of a possible meltdown for pension financing is highlighted by the OECD's [2004, p. 237] recent estimation that pension fund assets account for about 50% of member countries' market capitalization. For Porterba [1998], however, finding robust evidence in time series data for demography-driven changes in equilibrium returns on financial assets is difficult and he suggests caution in projecting large future changes in asset values. Yet, Lim and Weil [2003] show that in the presence of sufficiently large installation costs for physical capital, demographics do have the power to affect stock prices. Furthermore, hardly any study mentions or addresses the overall populations decline. However, the outlook to have fewer consumers in general is also unlikely to

[13]See Holzmann and Hinz [2005, p. 61-63]. Holzmann [2000] provides an assessment of the historic correlation of the various pillars' returns. The prospects on reform of the U.S. Social Security System has revitalized the discussion: see Lindbeck and Persson [2003] Shiller [2003*b*], Nataraj and Shoven [2003] or Modigliani and Muralidhar [2004]. Note furthermore that occupational pensions do not represent a pillar per se, but constitute merely a convenient implementation form for the mandatory saving accounts in the second pillar.

have a positive meanings for corporate earnings and asset prices.[14]
To sum up, both the funded as well as the unfunded financing method for pension systems seem to be profoundly affected by the demographic change: the former due to possible negative effects on financial returns and the latter by worsening old-age dependency ratios commanding ever higher contributions from wages. Consequently, the World Bank is right to promote a multi-pillar approach for pensions. But the diversification strategy can only work if the uncertainties in the pension financing methods are not perfectly correlated. Hence, the exact role of this diversification must be assessed in more detail if Pensionomics wants to proceed towards an optimized pension portfolio.

1.3 Objective and Synopsis

This work focuses on the portfolio perspective of the two financing mechanisms in the multi-pillar pension system introduced above. Given the fact that all pillars are finally financed either by funded savings or by a pay-as-you-go mechanism, the question arises how much each should contribute – facing the risks and uncertainties just described.

Portfolio Perspective

With the big challenges population aging poses on global societies, the aim is to develop a reference point for the combination of financing methods in old-age provisions. While this has some connection to health and long-term care, the predominant focus is on the pension subject. Given the World Bank's three roles for pension systems, the scope here is clearly limited to the consumption smoothing function. While redistribution interventions aiming at preventing poverty in old age are clearly justified from ethic grounds, they are difficult to theoretically integrate into an economic efficiency perspective. So is the insurance-like social safety net character of pension systems. The recent extension to five pillars, where redistribution is tax-financed, justifies this simplification: the politically desired level of redistribution is better achieved by a separate, tax-based program than in combination with the general unfunded pension system. Neither will the family, as the most important and fascinating organizational unit in societies and the World Bank's fifth pillar, play a prominent role. All this is not intended to trivialize the generational fairness discussions, but to confine the analysis to purely economic grounds

[14]Other contributions also assessing the impact of cohort sizes on asset markets are Bakshi and Chen [1994], Schieber and Shoven [1994], Yoo [1997], Bergantino [1998], Geanakoplos et al. [2002], or Davis and Li [2003]. Comments on Porterba's [2001] landmark article include Abel [2001] emphasizing the role of bequests and Campbell [2001] pointing out the rather constant aggregate saving rate. According to Porterba [2004] the current empirical findings are limited but suggest a modest impact of demography on asset values and returns.

with the traditional pillars of physical capital and human labor. Therefore, the portfolio perspective taken here is clearly not the exhaustive answer on the design of the multi-pillar pension problem, but only a subordinated and limited view providing a theoretic reference point. However, there are no restrictions concerning the goals which the World Bank set for pension schemes: life-time adequacy, affordability, sustainability and robustness are clear prerequisites for a genuine answer to the pension problem.
In this sense, this work focuses on the principle of *efficiency* in the management of risks. As the optimization of expected returns in relation to the associated risks through diversification of the elements is the key idea in any portfolio formation, so it is here for the design of the overall system: How and why shall funded and unfunded pension arrangements be combined in order to deliver retirement income more effectively and efficiently?

Overview

The answer to this question is neither a trivial nor a short one. It will be argued that the unfunded pay-as-you-go mechanism supplements funded pension provisions by allowing to virtually trade human capital. This insight is gained on basis of individual portfolio formation; yet, it also determines the role of each financing method in the multi-pillar approach. This portfolio understanding of the pension problem is built in three parts visualized in Fig. 1.1.

Part I comprises this Chap. 1's introduction to the phenomenon of global aging as motivation to analyze the pension problem as well as a survey of economic and financial theories related to it in Chapter 2. This establishes the theoretical requirements for the portfolio perspective and serves as methodological foundation for the analytical framework. This model is developed in Part II: Chapter 3 develops a general equilibrium model with overlapping generations that allows to address a hypothetical pension portfolio consisting of physical and human capital securities. This serves as first-best benchmark model for the replication following in Chap. 4. In this second-best approach, human capital is no longer marketable but a pay-as-you-go pension system and a consumption tax allow to achieve the same allocations as in the benchmark. Based on this indirect portfolio model, Part III derives implications for the pension problem. Chapter 5 analyzes the replication framework with respect to the requirements and to to its theoretical consistency. Important implications on the potentials of risk diversification and on the design of PAYGO schemes are derived. Chapter 6 concludes with a summary on these insights and an outlook on further research.

PART I: THE PROBLEM

- Aging as global phenomenon
- Pension problem arising

- Requirements for approach
- Economic and financial theory

Chap. 1
Global Aging
and Pensionomics

Chap. 2
Methodical
Foundation

PART II: THE MODEL

- Hypothetical portfolio construction
- First-best benchmark model

- Pay-as-you-go pension scheme
- Second-best replication

Chap. 3
Tradable
Human Capital

Chap. 4
Replication
with PAYGO

PART II: THE IMPLICATIONS

- Consistency and risk diversification
- Design of PAYGO schemes

- Insights on Pensionomics
- Outlook and further research

Chap. 5
Discussion
and Assessment

Chap. 6
Summary
and Conclusion

Fig. 1.2. General overview

Remarks: This thesis consists of three parts: The first one nests the pension problem into the context of global aging. Part two develops a second-best framework showing that a PAYGO system allows to replicate tradable human capital. The last part analyzes this model and derives conclusions for the pension problem.

2

Methodical Foundation

In fact, virtually nothing in the daily news will change how we live and what we do more than the global population shift now under way.

Kotlikoff and Burns [2004]: "The Coming Generational Storm", p. 35

With the understanding on the profoundness of the upcoming demographic change this chapter sets the requirements for the answer to the pension problem: without contradicting macroeconomic feedback mechanisms, the approach must simultaneously consider the funded and unfunded financing method in a financial risk-return perspective. To achieve this, Pensionomics will combine and extend prominent theoretical concepts from finance and economics. This review links these contribution to the pension issue and evaluates their limitations.

2.1 Outline and Requirements

Motivated by the deep impact of global aging, one must also build an understanding of the need for an extended economic model to address the portfolio perspective on pensions. This chapter realizes this in three major steps.
Firstly, based on the World Bank's intention to establish the multi-pillar approach as an effective and efficient way to deliver retirement income, the remainder of this section will develop three dimensions that interpret the bank's goals of adequacy, affordability, sustainability and robustness in an integrated manner. Secondly, Sect. 2.2 reviews major contributions of finance and economics with regard to these dimensions. As this step into Pensionomics builds on the neoclassical paradigm, the established findings on the portfolio approach, asset pricing, q-measure theory, economic growth and public finance form the theoretical foundation. The section also illustrates the linkages between the different theoretical areas. Though their rather separate development allowed for a certain division of labor between researchers, they have all some connection to the pension problem, but cannot cope with it completely. Based on these major building blocks, Chaps. 3 and 4 will therefore develop a distinctive replication framework for addressing the portfolio perspective on the pension problem. Because this follows a rather non-standard strategy,

Sect. 2.3 serves, finally, as a survey on the proposed replication approach and relates it to the specified dimensions.

Motivation of Requirements

Even though the pension issue seems to be a predominantly financial topic, it cannot be comprehensively solved within the financial perspective only. Since demography affects all economic markets, Pensionomics must not fall short of accurately addressing these changes and their consequences for retirement income. Principally, any analysis must thus include not only the market for capital assets but also for human labor and consumption goods as well as those for foreign exchange and the domestic currency. Two simplifying restrictions are made in this regard: this approach will limit itself to the analysis of a closed economy and thus abstract from international aspects in any of these markets and ignore the foreign exchange market. This is a clear simplification and is only tentatively justified by the fact that aging is a global phenomenon. Furthermore, there will also be no role for monetary aspects. In other words, the framework will be in real terms without a money market. The justification for this simplification is the very long-term focus of the analysis, as the demographic change realizes itself only gradually.[1]

Consistency in the remaining markets is a prerequisite for the approach. Economic growth theory, analyzing the effect of changes in a macroeconomic production factor like the population, is the natural link for them. Figure 2.1 illustrates the line of reasoning for the impact of demographic change.

On the one hand, population aging due to declining birth rates and higher life expectancy leads to decreasing cohort sizes. On the other hand, a different demographic composition of the total population implies a different aggregate savings behavior. Along these lines, demographic change affects human labor as well as physical capital. Both serve as input factors for the macroeconomic production of output, so that the goods market is affected, too. Based on the neoclassical marginal productivity theory, the distribution of income to labor and physical capital is thus also influenced by the altering demography. While physical capital's income reflects directly the uncertain return of funded pension savings, income from human labor is crucially determining the risk and return characteristics of the unfunded PAYGO system. The portfolio of both

[1] There is still surprisingly little research in the monetary field; see e.g. Auerbach and Herrmann [2002], Fitoussi [2003] or Davis [2004]. The international perspective is dominated by estimations of demography-driven capital flows like Börsch-Supan et al. [2002], Börsch-Supan et al. [2003], Brooks [2003], Domeij and Floden [2003], Feroli [2003], Börsch-Supan et al. [2005], or Domeij and Floden [2005]. Still, exchange rate effects are hardly assessed. Migration – as the labor market consequence of openness – is discussed, for instance, in Winkelmann and Zimmermann [1992], Alvarado and Creedy [1998], United Nations [2000] or Börsch-Supan [2003]. As indicated in Chap. 1, immigration turns out to be an unlikely cushion for older countries, as it requires substantially large magnitudes to offset the birth-decline; for *global* aging it can definitely not help.

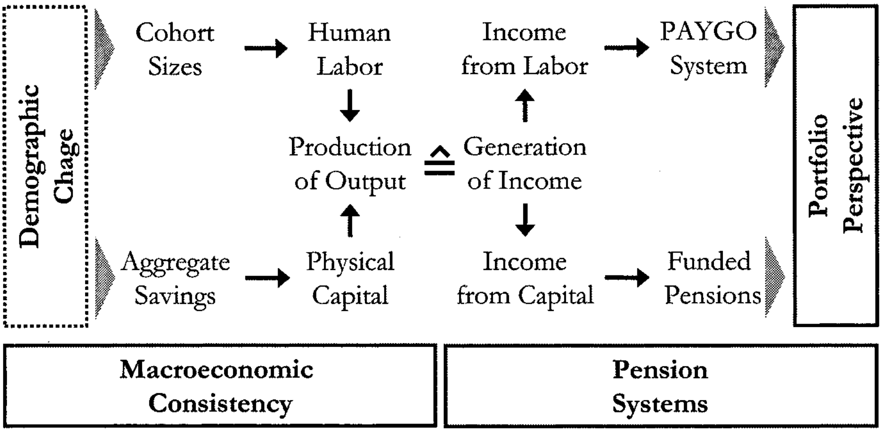

Fig. 2.1. Required dimensions

Remarks: A neoclassical intuition on the economy's interrelations implies that a genuine answer to the pension problem must meet three dimensions: (1) address funded and unfunded pension systems, (2) in a portfolio perspective under the condition of (3) macroeconomic consistency.

financing mechanisms reflects thus not only the demography-driven change in the capital market but also the aging's impact on the labor and the output market.

Required Dimensions

Based on this intuition, requirements in three dimensions can be established, which the investigation of the portfolio problem must satisfy. These dimensions are clearly related to the World Bank's goals of life-time adequacy, affordability, sustainability and robustness for the multi-pillar approach.
Firstly, the investigation must acknowledge that there are two basic forms of pension financing: the pay-as-you-go mechanism and funded savings. As mentioned earlier, this analysis will neither focus on the redistributional role nor on the institutional implementation. Therefore, the funded and the unfunded pension mechanisms constitute this analysis's understanding of *pensions systems* – even though the World Bank states five pillars.[2] As indicated in Fig. 2.1, the pay-as-you-go system is typically associated with the labor input factor and its income, whereas funded savings relate to investments and hence to the stock of physical capital. Since it will be required that the two

[2]It is therefore justified to equate the PAYGO pension system with public or unfunded pensions and the funded component with private savings. However, the replication framework in Chaps. 3 and 4 will suspend this distinction to some degree for methodological purposes.

systems must operate within their factor's financing capacity and allow individuals' participation in both, the World Bank's goal of individually and socially affordable systems is attained.
Secondly, the investigation must by consistent in a macroeconomic sense. The usual partial portfolio analysis in finance is based on the assumptions of marginality, i.e. the individual behavior has no impact on the aggregate development. The pension issue, however, requires to examine the economic activity of entire cohorts, whose aggregate behavior is not marginal but will have important consequences for the economic development and thus for the functioning of both pension systems. Therefore, *macroeconomic consistency* is mandatory: the implications of the analysis must still hold when all individual agents act as the portfolio view implies. In this sense, any solution must be subject to the condition of sustainable policies in the pay-as-you-go system and any other public intervention, above all concerning the expected population decline.
And thirdly, the investigation must – of course – take the *portfolio perspective.* As all pension systems are exposed to multiple risks in form of economic, demographic, and political uncertainty the World Bank's multi-pillar approach encourages diversification across different systems. This requires a finance-driven portfolio model to simultaneously address the risk-return characteristics of the funded and unfunded pension systems. This achieves robustness with respect to shocks from the mentioned uncertainty factors. Furthermore, a portfolio approach extended by the intertemporal consumption-savings decision – as usual in finance – addresses the adequacy of retirement income in terms of consumption smoothing over the individual lifetime.
These three dimensions are clear prerequisites for a genuine answer to the pension problem. Disregard of one will substantially alter the result, because funded and unfunded pension provisions co-exist, are characterized by specific risks and have feedback mechanism in an aggregate perspective.

Existing Research

Existing research neglects at least one of the dimensions – or does not explicitly address all of them. Though simulation studies like Brooks [2002] or Börsch-Supan et al. [2003] address the PAYGO system and macroeconomic aspect at detailed levels concerning the demographic composition and the public pension system, they ignore the portfolio dimension in combining the pensions mechanisms. Contributions from financial economics like Merton [1969*b*], Samuelson [1969] or Campbell and Viceira [2001*a*] focus thoroughly on the uncertainty aspect concerning allocations of savings or address uncertainty in the macroeconomic production relations – like Brock and Mirman [1972], Merton [1975] and Cox, Ingersoll and Ross [1985*a*,*b*]. Yet, they all ignore human labor as an input factor to production and hence the unfunded pension system as an alternative. Pure portfolio considerations of both systems like Dutta et al. [2000], Baxter and King [2001], Schacht [2001], Matsen and Thogersen [2004] or Borgmann [2005, Chap. 7] lack macroeconomic consistency. Given the remarkable future growth in the old-age dependency ratio as seen in Fig. 1.1 and the deep impact of global aging on macroeconomic development, portfolio-optimization in the spirit of Markowitz [1952, 1991] on basis of historical statistics must fall short of giving a genuine answer to the pension problem. Due to the currently high deficits of PAYGO systems, future returns are unlikely to be as favorable as past ones. Furthermore, the marginality assumption of the traditional financial portfolio approach is clearly violated when the savings-behavior of entire cohorts is considered.

2.2 Theoretical Background

This shortfall, however, does not imply that economic science cannot help with the pension problem. Established theories from finance and economics must rather be recombined to build Pensionomics and to shift the focus from scarcity of physical capital to scarcity of human beings. Past research has been fruitful in producing powerful concepts in the understanding of economic phenomena. Though the purpose of this section is not the history of these propositions, some historical remarks are appropriate given the interconnections of population history and the evolution of economics. Furthermore, it will be seen, that the extent of the pension problem requires Pensionomics to reunite different branches of economic thought, after having developed rather separately.

2.2.1 Neoclassical Paradigm

For combining and addressing the three dimensions of Fig. 2.1, the neoclassical paradigm with its "homo economicus" is a natural first reference point. With the concern of equilibrium and savings-determined investments, the schools of thought summarized under this term have been successful in producing approaches for a variety of economic issues – above all in the area of finance. Starting with the Marginalist Revolution Jevons [1871], Walras [1874], Marshall [1890], Clark [1899], Pareto [1906], and many others revived Bernoulli's [1738] idea of diminishing marginal utility and extended it to productivity. The reinterpretation of the former as expected utility by von Neumann and Morgenstern [1944] is doubtlessly one of the neoclassical centerpieces. The Paretian revival around Hicks [1939] and Samuelson [1947] was an important step towards the mathematization of mainstream economics like, for instance, Arrow and Debreu [1954]. The great variety of neoclassical thinkers ranges on the continental side from the Lausanne, Austrian, and Swedish Schools, to the Vienna Colloquium and on the Anglo-American side from the apologists, marginalists, and monetarists to the new classicals, the Cowles Comissions and the Chicago School. The landmark contribution of Fisher [1930] has proven to be of utmost importance for the course of development in financial and macroeconomic research.[3]

[3] See Henry [1990] for a detailed history of the neoclassical paradigm.

Fisher Separation

Fisher [1930] introduced the Austrian intertemporal theories to the American neoclassics; Hirshleifer [1958, 1970] helped to revitalize and interpret his idea of the capital market. The two central results are known as the *Fisher Separation Theorem*: the firm's investment decision is independent of the preferences of the owner and the investment decision is independent of the financing decision. Figure 2.2 illustrates this two-stage budgeting process.

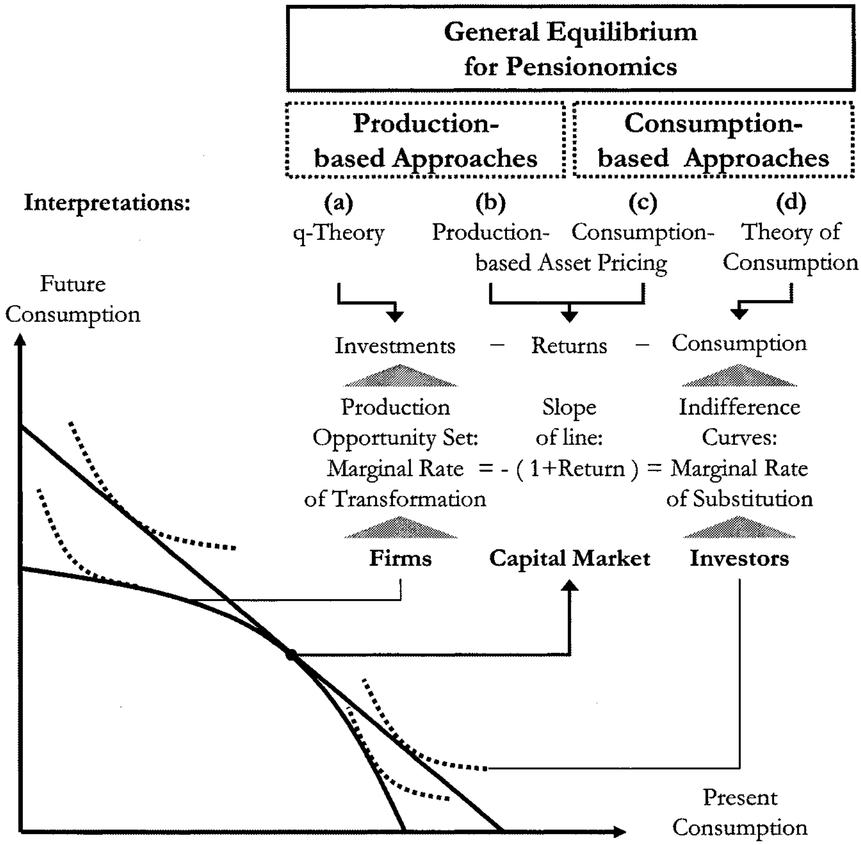

Fig. 2.2. Fisher separation and interpretations

Remarks: The Fisher Separation as central result of a two-stage intertemporal optimization with a perfect capital market asserts that the firms' objective is the maximization of their present value, regardless of the investors' preferences. It allowed for the independent development of four different interpretations, (a) to (d), all to be addressed in this step into Pensionomics.

On the one hand, utility-maximizing investors, characterized by indifference curves, trade off current and future consumption. On the other hand, profit-maximizing firms adjust their investment input. This productive investment allows to transform forgone present consumption into future consumable units. The shapes of the production opportunity set and indifference curves reflects the Marginalists' core concept of diminishing marginal productivity and utility. The importance of a perfect capital market in this intertemporal optimization must not be underestimated: it allows an efficient transfer of funds between borrowers and lenders. For investors as well as for firms, the return on the capital market is the *correct* measure of opportunity costs. This allows to equalize the marginal rates of substitution and transformation. The firms can maximize their present value by realizing the production optimal point independent of the investors' preferences, so that all investment opportunities yielding a higher return than the capital market are realized. The investors, on the other hand, meet their consumption preferences by borrowing or lending within the capital market. At the prevailing equilibrium interest rate, borrowing and lending nets exactly to the financing requirements of the firms.[4]

The Fisher Separation Theorem must be considered the microeconomic foundation for a large number of concepts in the areas of finance and macroeconomic theory. Above all, it led to the relatively independent development of two major branches and four interpretations, which are also indicated in Fig. 2.2: the production-based and the consumption-based approach. The former is based on the producer's first order conditions for optimal intertemporal investment demand based on a production function. This production-based approach focuses thus on the marginal rate of intertemporal transformation and is interested in the process of investments and returns. Using the return process for predictions on investment, it is an interpretation of (a) the q-Theory of Tobin [1969]. Fixing or modeling the return process one has (b) a production-based asset pricing model for returns like Cox, Ingersoll and Ross [1985*a*,*b*]. The logic of the latter branch is analogous, yet based on the consumer's first order conditions for intertemporal consumption derived from the utility rationale. It focuses on the marginal rate of intertemporal consumption substitution and the joint evolution of consumption and returns. Fixing the consumption process and interfering predictions about returns one has (c) a consumption-based asset pricing model like Breeden [1979]; taking the return process as given, it is (d) a theory of consumption, like for instance Friedman [1957].[5]

Though this methodical split has been extremely fruitful for the development

---

[4]See Fama and Miller [1972] and Copeland and Weston [1988] for an extensive discussion of the neoclassical foundation of finance. The interest rates' linkage to production-based economics is nicely reflected in the fact that the typically notation for them is not "i" but "r" – for *rental* of physical capital; see Wheelan [2002, p. 121] for an anecdotic illustration.

[5]This distinction of two approaches and four interpretations draws on Cochrane [1991, 2005].

of partial models, Pensionomics requires a general equilibrium approach consistently addressing marginal rates of substitution *and* transformation. To understand the connections of the four interpretations to the pension issue, they are presented in the following.

Utility Concept and Consumption Smoothing

Starting with the last interpretation, (d), it is noteworthy that the feature of consumption smoothing is a common pattern of financial asset-pricing models and economic hypotheses of income. After Duesenberry's [1949] first attempt to improve the simple Keynesian consumption relation, the proposition of Friedman's [1957] "Permanent Income Hypothesis" and Modigliani and Brumberg's [1954] "Life Cycle Hypothesis" – with more explicit emphasis on age – were major steps towards a microeconomic-founded understanding of consumption behavior. Both distinguish life-time wealth from transitory changes in income and argue that consumption depends on wealth rather than on current income. In an extension of Fisher's [1930] approach consumers are faced with an expected income stream and – based on an intertemporal optimization – allocate their consumption over the lifetime by means of the capital market.[6]

This optimization is the typical neoclassical rationale of maximizing utility from consumption. To put it formally, persons are modeled as rationale agents with von Neumann-Morgenstern preferences and they maximize their expected utility as given by the time separable intertemporal utility function

$$U(t) = \mathbb{E}_t \left[\sum_{i=0}^{\infty} \beta^i \times u_{t+i}\left(c(t+i)\right) \right], \tag{2.2.1}$$

where the instantaneous utility functions $u_t(c(t))$ are monotone increasing and strictly concave in consumption $c(t)$ at time t. β is the subjective discount factor.[7] This maximization is constraint by the intertemporal budget constraint of the agent's wealth, $w(t)$,

$$\tilde{w}(t+1) = (w(t) - c(t))\, \tilde{R}_T(t+1), \tag{2.2.2}$$

according to which investing the unconsumed portion of wealth yields a total gross return of $\tilde{R}_T(t+1)$. The idea of wealth, $w(t)$, is thus the *central* concept for linking the streams of consumption and income. Ignoring bequests, all wealth is derived from income – from human labor and physical capital – and serves sooner or later the agents' ultimate goal of consumption:[8]

---

[6] See Modigliani [1986] for a summary on the life-cycle consumption approaches.

[7] For more detailed expositions of utility theory see Ingersoll [1987] or Gollier [2001]. Campbell and Viceira [2001*a*, Chapt. 2] provide an introduction to the utility concept underlying portfolio theory. The seminal work of Ramsey [1928] seems to be the first analysis of the multiperiod consumption problem.

[8] Note that bequests can easily be addressed by viewing them as initial endowment income or like a terminal consumption similar to Merton's [1969*b*] bequest term.

$$\mathbb{E}_t \left[\sum_{i=0}^{\infty} \frac{c(t+i)}{\prod_{j=0}^{i} \tilde{R}_T(t+j)} \right] = w(t) = \mathbb{E}_t \left[\sum_{i=0}^{\infty} \frac{y_T(t+i)}{\prod_{j=0}^{i} \tilde{R}_T(t+j)} \right] . \quad (2.2.3)$$

With the left hand side of (2.2.3) Campbell [1993] gives this explicit present value budget constraint interpretation. The macroeconomic hypotheses of income, on the other hand, emphasize the link of future total income streams, $y_T(t)$, and wealth as shosn on above's right hand side. When the current stream of income is unexpectedly high in relation to overall wealth, investors will smooth consumption into the future by additional savings.

The Life Cycle Hypothesis addresses the finiteness of an agent's life and the fact of retirement explicitly. For (2.2.1) and (2.2.3) this implies that the summations are not to infinity but end at a certain terminal year and the income stream from working terminates at some retirement age. The main result of the life-cycle theory is the hump-shaped savings profile over the life-cycle: to compensate the lower income during retirement and to avoid a drop in utility at the point of retirement, agents save some fraction of their income during their working life and dissave during retirement. Though empirical tests of this prediction using actual household behavior largely failed, the hypothesis of Modigliani and Brumberg [1954] and Modigliani and Ando [1963] became the main theory of aggregate savings and stimulated a large quantity of research. Noteworthy is the fact that intergenerational transfers, like those of an unfunded pension system, are one explanation for the empirical failure as for instance that of Kotlikoff and Summers [1981]. These transfers distort or overcompensate the pure life-cycle effects and must thus be thoroughly addressed. Still, regarding asset allocation and portfolio choice the Life Cycle Hypothesis is the cornerstone of many practical financial investment strategies and receives increasing academic attention.

The agents' life-cycle with its distinction of working period and retirement and the utility-driven feature of consumption smoothing must be considered as the origin of any pension idea and will thus be the backbone of the framework. However, the concepts discussed here neither address the macroeconomic background nor do they include the portfolio perspective.

2.2.2 Portfolio Approach

While the diversification idea has already been introduced in the description of the multi-pillar pension strategy, the portfolio approach is normally considered in the context of agents' investment decisions by allocating wealth among a variety of available securities. Markowitz [1952] and Roy [1952] are the first contributions to clearly formulate and solve this problem of portfolio selection and hence mark the beginning of modern portfolio theory. In particular Markowitz [1952, 1991] and Tobin [1958] use the axiomatization of expected utility by von Neumann and Morgenstern [1944] to address the decision under uncertainty, which the portfolio selection involves. While the

former contributions may be viewed microeconomic and establish the well-known rule of expected mean returns and variance of returns, Tobin [1958] addresses a macroeconomic problem and derives the important Two-fund Separation.[9]
Key to any diversification is the idea of agents' risk aversion as implied by the concavity of the utility function in (2.2.1): a low level of future consumption has a more severe negative effect on utility than an equally probable high consumption has a positive one. Based on this curvature of the utility function relative to its slope, Arrow [1971] and Pratt [1964] established specific definitions of risk aversion – in case of the absolute risk aversion dependent on the given level of wealth and independent of it in case of the relative one. This initial analysis had actually been based on utility functions of wealth. It is now well understood that the limitation to mean-variance analysis is justified by the assumption of quadratic utility, normally distributed asset returns and exponential utility, or lognormally distributed returns with power utility. Cass and Stiglitz [1970] have shown that the Two-fund Separation holds when preferences are either quadratic or of the constant relative risk aversion class.[10]

Consumption and Portfolio Selection

Since the mean-variance portfolio theory as just described scrutinizes on the asset selection problem for investments of one period only, it is sufficient that investors care about the distribution of wealth at the end of that period. However, this is theoretically and empirically troublesome, because agents are concerned with the standard of living affordable and not with wealth for its own sake. Equations (2.2.1) to (2.2.3) reflect this, as they focus on the *consumption* stream that the level of wealth does support over the life-time. As the restrictive assumption of single-period investments was relaxed, major progress in portfolio theory produced two principal lines: the discrete time multiperiod models like Samuelson [1969], Fama [1970*b*], or Hakansson [1970], and their continuous-time counterparts as Merton [1969*b*, 1971, 1973], Breeden [1979, 1986], and others. Multiperiod problems introduce the possibility of stochastic changes in the opportunity set. Above all, Merton's seminal contributions achieved the extension of the intertemporal consumption-investment problem and showed how agents trade off present and future consumption and that their portfolio allocation is heavily affected by the possibility of time-variation in returns. A Three-Fund Portfolio Separation is obtained, in

[9] See Fama [1976], Ingersoll [1987, Chaps. 3 and 4], Copeland and Weston [1988, Chap. 4], or Constantinides and Malliaris [1995] for analytical summaries of portfolio theory. Markowitz [1999] gives a historical overview.

[10] See Campbell and Viceira [2001*a*, Chap. 2]. It is noteworthy that also the macroeconomic theories of income and consumption address the aspect of risk aversion with the idea of precautionary savings as proposed by Leland [1968], Sandmo [1970], and Kimball [1990].

which a hedging component with respect to the investment opportunity set complements the mean-variance case.[11]

Labor Income and Human Capital

Even though these classical results focus on life-time portfolio selection they are only partially useful for the pension problem. This is because they ignore labor income, which is typically the most important source of income over the life cycle. Instead, the understanding of wealth in most contributions of financial economics is implicitly limited to a physical one: wealth is seen as the present value of future claims derived from the ownership of real capital. However, as the neoclassical intuition of Fig. 2.1 indicated, human labor contributes not only equally to the generation of output, but also drives the performance of the unfunded pension pillar. Since it does not matter for the individual agent from which source of income his wealth is generated as long as it allows for consumption, the understanding of wealth here does not exclude any sources of income. Equation (2.2.3) reflects this by the fact that $y_T(t)$ comprises income from physical capital, i.e. financial wealth in the classical sense, as well as income from human labor. Therefore, it is essential to thoroughly include labor income in the portfolio approach.
Like the claims derived from ownership of physical capital, future labor income can be associated with a specific present value and constitutes in this sense "human wealth" or – in the spirit of Schultz [1961] and Becker [1962] – "human capital". Still, this analogy between human capital and physical one breaks down in one important respect: while property rights over ordinary, inanimate physical capital are typically easily transferable by sale from one owner to another, human capital is by definition *inseparably* embedded in the individual human being. This creates a standard moral hazard problem: contracts committing an agent to work are not legally enforceable as this would correspond to a form of slavery. As buyers realize that an agent has no incentive to work after selling his capitalized labor income, this market collapses and human capital becomes virtually a non-tradable asset. Only its temporary labor services can be directly bought and sold on the market, whereas markets for physical capital goods do function easily and smoothly.[12]

---

[11] Merton [1990] or Duffie [2001] provide general treatments of portfolio choice in continuous time. Focussing on the aspect of asset allocation with predictability of returns Campbell and Viceira [1999, 2001*b*], Barberis [2000], Campbell, Chan and Viceira [2003]; Campbell, Viceira and White [2003] or Campbell et al. [2004] exploit the vector-auto-regression procedure and develop several approximate solutions, since the exact analytical ones are usually not obtainable. Campbell and Viceira [2001*a*] is a textbook treatment of this.

[12] Furthermore agents can influence the risk-return characteristics of their human capital by varying how much they work and by education. As mentioned earlier, this aspect of human capital *formation*, which dominates the analysis of Becker [1993], is not addressed in this framework. Admittedly, human capital is to some minor

For the portfolio problem this particularity of labor income and human capital drastically complicates the analysis: on the one hand, income from labor is risky and represents a significant, undiversifiable portion of life-time wealth so that its risk-return characteristics should play a dominant role in the determination of the agent's asset allocation. On the other hand, the particularities of human capital impede the application of financial economics' standard techniques. Core of the problem is a non-linearity problem, which the non-tradability of human capital introduces into the budget constraint. If nonmarketable, labor income *cannot* be summarized into the definition of wealth as given in (2.2.3). The indicated discounting – at the risk-adequate rate – is based on the Fisher Separation as explained in Sect. 2.2.1 and relies thus centrally on the idea of a *tradable* claim. Instead of this, the current flow of *labor* income, $y_\omega(t+1)$, must be added additively to the gross return $\tilde{R}_W(t+1)$ of tradable wealth in each period:

$$\tilde{w}(t+1) = (w(t) - c(t))\,\tilde{R}_W(t+1) + \tilde{y}_\omega(t+1)$$
$$\text{or alternatively}\quad \tilde{w}(t+1) = (w(t) + y_\omega(t) - c(t))\,\tilde{R}_W(t+1). \qquad (2.2.4)$$

The additional term causes a *breakdown* of the usual connection of wealth as a stock variable and consumption as a flow variable, since labor income is another flow component, that is not completely controllable. Therefore, the simultaneous addressing of returns on investments in physical capital and labor income as required for the pension problem causes serious problems for the standard portfolio theory. Contributions like Williams [1978], Heaton and Lucas [1997, 2000], Viceira [2001], or Campbell and Viceira [2001*a*, Chap. 6] have achieved some integrations, but rely on simplifications. They focus on single or two-period problems, apply special utility or asset return assumptions, use approximation procedures or rely heavily on numerical procedures to solve models with liquid and illiquid components of wealth. Equally dependent on numerical approaches are the studies testing and estimating the Life Cycle Hypothesis' implications for portfolio selection, like Bodie et al. [1992], Bakshi and Chen [1994], Campbell et al. [2001], Campbell and Viceira [2001*a*, Chap. 7], Gomes and Michaelides [2004] or Campbell et al. [2005]. Ameriks and Zeldes [2001] highlight the unsolvable identification problem in this regard: because time, age and birth year are linearly dependent on each other effects driven by them *cannot* be identified separately without further assumptions. This problem seriously impairs any empirical study in this field.[13]

extend tradable in cases of borrowing against future labor income as collateral or specific cases like book publishing contracts etc.; see Fama and Schwert [1977]. Yet, given the aggregate, economy-wide perspective required for the pension problem these exceptions are negligible.

[13]Fixed commitments like housing are often ignored, even though they represent a significant portion of wealth and deliver an important consumption part. Cocco [2005] is one of the few attempts to include it in the life-cycle portfolio model. Corresponding to Fig. 2.1's focus on only two macroeconomic input factors, namely

For the step into Pensionomics it is essential to address the individual agent's simultaneous holding of tradable financial wealth as well as the changes in his human capital in course of the life cycle. Furthermore, unfunded pension systems introduce an additional mean for intertemporal consumption smoothing. The corresponding claims can be considered as another form of wealth – namely transfer wealth as introduced by Feldstein [1974], Kotlikoff and Summers [1981], Modigliani [1988], and Lee [1994*b*]. While earlier research has often treated this as an implicit holding of a *riskless* asset, this assumption is hardly maintainable given the vast future demographic changes discussed in Chap. 1. It is thus central to the diversification within the multi-pillar approach to integrate human capital into the portfolio perspective. Understanding the diversification idea as a logical extension to the objective of intertemporal consumption smoothing, this concludes the description of consumption interpretation (d) in Fig. 2.2.

2.2.3 Asset Pricing

While so far the return process has been take as given, the asset pricing interpretations reverse the question and try to explain the observed process of prices and the implied asset returns. The explanatory input is the process of investments in case of Fig. 2.2's interpretation (b), and the observed consumption process in case of interpretation (c).

Consumption-based Asset Pricing

Continuing in the consumption-based field, the research question can be understood as the identification of pricing functions that map future asset payoffs into current prices. Thereby the stochastic discount factor or "SDF"-concept has evolved as the dominant representation. While its origins lie in the Arrow–Debreu model and its application to option pricing by Cox and Ross [1976] and Ross [1978] along Ross's [1976] Arbitrage Pricing Theory, Harrison and Kreps [1979] provided a definitive theoretical treatment in continuous time. In discrete time, the SDF methodology has substantially been pioneered by Rubinstein [1976], Lucas [1978], Grossman and Shiller [1981], Hansen and Richard [1981], Hansen and Jagannathan [1991], and Cochrane and Hansen [1992].

Based on the rationale of (2.2.1), the SDF approach focuses on a basic pricing equation related to Fisher's [1930] approach:[14] if the agent can freely trade his wealth by purchasing an extra unit of an asset A at price $p_A(t)$ and collecting its future gross payoff $\tilde{X}_A(t+1)$, he will optimally trade off marginal

physical capital and human labor, land as a potential third one as well as real property investments will be ignored in this framework.

[14] Appendix A.1 provides an intuitive interpretation of the SDF as state-dependent Fisher Separation, where an asset's price can be written as its probability weighted average payoff.

opportunity costs of consumption and the expected marginal benefit of this investment. Time-separability of the utility function $U(t)$ allows to formulate the first-order condition of the agents' intertemporal optimization in a stochastic Euler equation as

$$\frac{\partial U(t)}{\partial c(t)} \times p_A(t) = \mathbb{E}_t \left[\beta \frac{\partial U(t)}{\partial c(t+1)} \times \tilde{X}_A(t+1) \right] . \qquad (2.2.5)$$

Defining now the stochastic discount factor as

$$\tilde{M}(t+1) \equiv \frac{\frac{\partial U(t)}{\partial \tilde{c}(t+1)}}{\frac{\partial U(t)}{\partial c(t)}} = \beta \frac{\frac{\partial u_{t+1}}{\partial \tilde{c}(t+1)}}{\frac{\partial u_t}{\partial c(t)}} \qquad (2.2.6)$$

and dividing (2.2.5) on both sides by $p_A(t)$, one has the basic pricing equation as condition on the asset's gross return $\tilde{R}_A(t+1) \boxminus \tilde{X}_A(t+1)/p_A(t)$:

$$1 = \mathbb{E}_t \left[\tilde{M}(t+1) \times \tilde{R}_A(t+1) \right] . \qquad (2.2.7)$$

Equation (2.2.7) is *the* central relation of financial economics. Though the contribution of Breeden [1979] offers one of the clearest links to consumption, the problem is that consumption data is poorly measured in practice. Hence, the manifold models in the theory of finance can be understood as usage of other determinants of consumption and as a direct modeling of marginal utility in terms of these factors: wealth in case of the Capital Asset Pricing Model of Sharpe [1964], Lintner [1965], and Mossin [1966], innovations about future opportunities in case of Merton's [1973] Intertemporal Capital Asset Pricing Model, or other purely statistical factor loadings in case of Ross's [1976] Arbitrage Pricing Theory and its intertemporal version like in Stambaugh [1983]. Furthermore, the mere existence of the representation (2.2.7) and the fact that marginal utility is positive, this equation is the basis for the arbitrage or near-arbitrage pricing techniques. Underlying principle is always (2.2.7) or manipulations and specializations of it, which allow for implications on returns, excess returns, riskfree returns, payoffs of zero-cost portfolios etc. Being an optimality condition (2.2.7) must hold for any asset.[15]

Endowment Economy

From an aggregated perspective, the similarity of the SDF-concept to the Fisher Separation is evident, as the SDF is inherently related to the agents' marginal rate of substitution and the returns reflect the economy's marginal

[15] See Campbell [2000] or Campbell [2003*a*] for summaries of the SDF approach. Ingersoll [1987] and Duffie [2001] provide textbook treatments. Cochrane [2001] restates the whole consumption-based asset pricing theory within this framework and provides the relation to the econometrics, which are scrutinized on in Campbell et al. [1997].

rate of transformation. Yet, aggregation from individual agents to representations of the entire economy is non-trivial.[16] Nevertheless, Fig. 2.2 indicates that (2.2.7) must be supplemented by a description of the intertemporal productive possibilities, in order to find asset prices and consumption in terms of truly exogenous variables. Before turning to this production-side, an important simplification must be addressed. The concept of an "endowment economy" as developed by Lucas [1978] relies on an extreme variant of the production technology.[17] It can be understood as a production opportunity set without intertemporal transformation, so that agents *cannot* convert present consumption goods into future ones by saving and investments. Instead, each period there is a certain endowment of goods, so that the central dot in Fig. 2.2 would characterize the production opportunity set. Since the capital market's returns, i.e. the asset prices, will still adjust until agents are happy consuming this endowment process, this approach allows to abstract from the production side. In analogy to the yielding of fruit by a tree, which can be picked and consumed each year, the endowment economy is also referred to as "Lucas fruit-tree model".

However, this endowment perspective has produced several empirical challenges of the asset pricing paradigm. Above all, the Equity Premium Puzzle of Mehra and Prescott [1985] illustrates a quantitative mismatch of theory and data, first identified by Shiller [1982]: historically, the excess return on equities over the riskfree return has been an order of magnitude greater than can be rationalized as a premium for bearing risk within the standard asset pricing paradigm. Other challenges include Weil's [1989*b*] puzzle of the riskfree rate and the equity volatility problem of Campbell [2003*a*]. The puzzles have resulted in a vast amount of literature producing various explanations to solve the mismatch: alternative utility specifications including behaviouralistic elements, statistical explanations or market imperfections, and others.[18]

From the Pensionomics' point of view the contributions focussing on life-cycle and generational aspects are especially interesting. Examples include Mankiw and Zeldes [1991], Storesletten et al. [2001, 2004], or Ang and Maddaloni [2003]. Financial research is discovering demographic elements as valuable components for the economic models. Above all, the contribution of Constantinides, Donaldson and Mehra [2002] addresses the particularities of an overlapping generations model, which will be introduced in more detail in Sect. 2.2.6. They explore the fact that the relative attractiveness of risky equity investments changes in course of the life-cycle: for younger agents stocks

[16] See for instance Rubinstein [1974], Constantinides [1980, 1982], or Scheinkman [1989].

[17] The Lucas [1978] landmark contribution also highlights the concept of rational expectations in the spirit of Muth [1961] and the proposition of efficiency in the capital market of Fama [1970*a*].

[18] See Kocherlakota [1996], Constantinides [2002], Jäkel [2002], Mehra [2003], or Campbell [2003*b*] for summaries on the empirical challenges, the Equity Premium Puzzle, and proposed solutions.

are a desirable asset, representing a hedge against fluctuations in the future labor income as long as stocks and wages are not highly correlated. The middle-aged agents, on the other hand, require a relatively high return for holding equities, since their labor income uncertainty is largely resolved and fluctuations in retirement income are only due to the equity risk. Consequently the equilibrium characteristics of equity depend on who the predominant holder is. The central element for this answer to the Equity Premium Puzzle is the assumption, that younger agents are constraint from borrowing against their future labor income and are thus effectively excluded from the stock market. Hence, middle-aged agents represent the marginal investor, which justifies a relatively high equity premium. This article is one of the major building blocks for Chaps. 3 and 4's replication framework.

Direct Production Models

Based on the intuition of Fig. 2.1, it is logical that the endowment economy simplification is hardly useful for addressing the simultaneous consequences of demographic change on human labor and physical capital. Instead, Pensionomics does require a genuine production side. The production-based asset pricing interpretation (b) in Fig. 2.2 does contribute to this. Analogous to its consumption counterpart, it allows for predictions about asset returns and prices – however, these are now based on the firms' stochastic intertemporal marginal rate of transformation. The specific form of the production technology underlying this transformation becomes central. Sundaresan [1984], Cox, Ingersoll and Ross [1985*a*,*b*], or Constantinides [1992] attempt to address the production side, but assume a *linear* technology, so that the marginal rate of transformation – and hence the return on real capital is *not* affected by the amount invested. This means that in Fig. 2.2 the firms' production opportunity set would coincide with the straight line characterizing the capital market. Given the neoclassic paradigm of diminishing marginal utility and productivity indicated in Sect. 2.2.1 this specification is also troublesome for Pensionomics' long-run perspective, i.e. it does not meet the requirement of macroeconomic consistency.[19]

A true step towards general equilibrium relies thus on a non-trivial idea of the production side and its implication for intertemporal transformation of consumable units. Generalizations in this regard include Brock [1979, 1982] or Campbell [1994] applying stochastic versions of the neoclassical growth model like Brock and Mirman [1972] or Malliaris and Brock [1982]. Rouwenhorst [1995] and Boldrin et al. [2001] focus on the asset pricing implications of the

[19] Section 2.2.5 will discuss the theoretical background of the production side more thoroughly and address an important aspect completely ignored so far: human labor as the second macroeconomic input factor. Note, however, that the assumption of constant returns to scale is less troublesome for the relatively short-run prediction of the production-based asset pricing models than it would be for the long-run focus of Pensionomics.

real business cycle model of Kydland and Prescott [1982], which is related to the stochastic growth model. Especially insightful for the Pensionomics framework to be developed is Balvers, Cosimano and McDonald [1990], who draw on both influences and develop the idea of a representative firm in analogy to the representative consumer. However, most of these direct production-based approaches tend to generate counterfactual implications for asset pricing: the problem is that while investments are rather smooth, asset returns are highly volatile. Furthermore, the decomposition of asset price variation into a direct technology shock effect and an indirect effect due to accumulation of physical capital weakens the link of investments and returns. Major advances are only achieved with the inclusion of Tobin's [1969] q-Theory.[20]

2.2.4 Production and q-Theory

Therefore, this section briefly introduces the q-measure of Tobin [1969], which represents interpretation (a) of Fig. 2.2 and returns then to its applications in the production-based interpretation (b).

Tobin's q and Adjustment Costs

In an empirical implementation of Keynes's [1936] notion that physical capital investment becomes more attractive as its value relative to the cost of acquiring it increases, Tobin [1969] – and Brainard and Tobin [1968] – introduced the q-measure as the ratio of a firm's market value to the replacement cost of its physical capital stock. Based on the neoclassical Jorgenson [1963]-model of investments, Mussa [1977], Hayashi [1982], and Abel [1983] linked the q-idea formally to the concept of adjustment costs, which had been introduced by the Eisner and Strotz [1963]-Lucas [1967]-Gould [1968] theory of investments and extended in Lucas and Prescott [1971].
Underlying is the contrast of Fisher's [1930] determination of interest rates by investment *flows* and Clark's [1899] standard neoclassical relation of interest rates to the productivity of the physical capital *stock*. The concept of adjustment costs – or alternatively gestation lags – and the interpretation as q-ratio allows to combine this: essentially, investment flows do not become installed physical capital directly, but must be converted at some adjustment costs. Hence, the marginal efficiency of additional investments corresponds to the marginal productivity of physical capital less the marginal adjustment costs. Since the neoclassic paradigm implies that the firms' valuations on the stock market reflect the productivity of the capital stock, the adjustment cost allow for a wedge between this price of installed real capital and the price of uninstalled consumable units on the output market, which serve as investment input. The ratio of these two is captured by Tobin's [1969] q-measure, which can thus be understood as the ratio of the marginal efficiency of investments to the interest rate.

[20] See Lettau [2003] or Cochrane [2005].

The argument can also be turned the other way around, yielding the q-Theory of investment, which is Fig. 2.2's interpretation (a): if the q-ratio is greater than one, a firm's market value – as implicit in its capital market returns – exceeds the value of the company's recorded assets. Such a high q-ratio encourages companies to invest more, since the additional installed physical capital is valued higher than the price they pay for the investments. If, on the other hand, the q-measure is less than unity, the capital market is effectively restraining the company from investing, because it would have to issue relatively cheap securities to finance it. Reversion of these tendencies is implied by convexity of the adjustment function, which follows the marginalists' tradition: the adjustment costs rise with the amount of investments. This causes a difference between the average q-ratio and the marginal one. Abel and Eberly [1994, 1996] relate these two measures in a stochastic setting and study the effect of reduced reversibility of investments. While in macroeconomic applications of the q-Theory – like Abel and Blanchard [1986] – this option-like character is barely addressed, on a microeconomic firm-level the real option approach has gained much attention.[21]

q-Theory in Asset Pricing

Though long neglected, the adjustment cost interpretation of Tobin's [1969] q-measure receives increasing attention in the asset pricing field. For the production-based interpretation (b) of Fig. 2.2, the wedge between the price of installed real capital and the price of consumable units offers additional flexibility: the adjustment costs attenuate the volatility implied by the return process and enable so smoother investment implications. But the q-Theory has still a clear focus on the firms' marginal rate of transformation. Major contributions that revived the literature in this direction are Restoy and Rockinger [1994], Cochrane [1991, 1996], Jermann [1998] and Lamont [2000]. With increasing success they incorporate the real business cycle model, respectively an one-sector stochastic growth model, and technologies with convex adjustment costs for the physical capital stock into asset-pricing models. As stochastic shocks cause the q-measure to divert from unity, these model do not suffer – like the direct production approaches – from the troublesome implication that consumable units and real capital can be instantaneously transferred into each other. The applications are now manifold. Because the book-to-market ratio is related to the inverse of the q-measure, interpretations like Li et al. [2005] and others shed new light on the multi-factor asset pricing literature of Fama and French [1989, 1992, 1993, 1995, 1996]. Multiple sector models like Gomes et al. [2003] or Zhang [2005] bring the asset pricing field more in line with the real business cycle literature. Also the irreversibility aspect is newly

[21] See Dixit and Pindyck [1994] for a survey and textbook treatment of the literature on irreversible investments and its relation to the q-Theory.

addressed, for instance in Kogan [2004].[22]
For Pensionomics the q-ratio is interesting because Lim and Weil [2003] illustrated that demographics command an influence on stock prices in presence of sufficiently large adjustment costs. Abel [2003] focuses also on this effect and addresses the consequences of a simultaneous public pension scheme – eventually with some funding. While he does not address the portfolio idea for combining the pension systems, and rather takes the public system as given, his contribution and its modeling of the adjustment costs will still be heavily used in Chaps. 3 and 4's replication framework. However, for the pension problem, an important aspect is still unaddressed: the consistent modeling of returns on physical capital *and* of wages for human labor service. While most of the asset pricing models focus exclusively on the former, even the prominent ones, which apply some labor income variables, like Jagannathan and Wang [1996] or Campbell [1996], must make a simplifying assumption: the return on human capital is approximated by random changes in future wages discounted to their present value at a proxy discount rate. The conceptual problem is evident in Fig. 2.2: underlying the approach of Fisher [1930] is a simplifying assumption of a production relation with physical capital as *only* input factor. Given the understanding of Fig. 2.1 and the Chap. 1's emphasis, that the demographic change will cause an unprecedented tendency towards scarcity of *human beings*, the statistical extrapolation of past labor income processes would be unhelpful. Instead, a more thorough inclusion of human labor is mandatory.

2.2.5 Growth Theory

Neoclassical growth theory does allow for such a consistent linkage of human labor and real capital in the long-run required for the pension problem. Introduced as the underlying intuition in Sect. 2.1, this theory does not only reflect the productive aspect of the two input factors for the economy, but gives also a rationale for the distribution of income to them. As Fig. 2.1 illustrates, it is thus a central concept for the set requirements of macroeconomic consistency and the simultaneous addressing of the PAYGO and funded pension system.

Economic Growth

The phenomenon of economic growth refers mainly to the observation that per capita output and income rise steadily over time. Growth theory attempts to understand this and the fact, that per capita incomes vary greatly across countries. With this focus economic growth must be considered as a core concern of the classical economists like Smith [1776], Ricardo [1817] or Malthus [1798], who provided many of the basic indigents for the theory. Furthermore, growth theory also strives to explain the observation, that the *level* of returns

[22] See Cochrane [2005] for a recent survey of models at these intersections between macroeconomics and finance.

to real capital – as reflected in interest rates and equity returns – has been constant in the long run, even though the capital stock per worker has substantially grown. This observation is complemented by the fact that the shares of labor and physical capital in national income are nearly constant: though fluctuating, approximately two thirds are paid to labor services and one third to real capital.[23] Despite the fact, that growth theory has been developed during the three-centuries old trend of increasing populations, its analytical framework has been formulated so that it can handle diminishing populations as well.

To address the stylized fact, the classicals focussed on the supply of consumable goods in an economy and developed the vital concept of a production function. The contrast of the classical proposition, that macroeconomic output is based on the productive usage of the physical capital *stock*, human labor and land, and Fisher's [1930] intertemporal optimization based on investment *flows* is again key to the discussion. In this sense the illustration of the Fisher Separation in Fig. 2.2 simplifies the classical production notion in two ways: First, this interpretation is based on the understanding that all capital is circulating one – in other words, that there is no difference between investments and the stock of real capital.[24] Second, by holding the amount of labor and land constant, these inputs are effectively eliminated from the consideration. While land will also be ignored in this step towards Pensionomics,[25] the abstraction from the labor input would be like ignoring the upcoming worsening in the old-age dependency ratio shown in Fig. 1.1. For the pension problem the underlying production relation must thus be of more general nature: it must address the capital augmenting effect of investments and it must allow for varying levels of labor. In this sense, one could interpret the production opportunity set in Fig. 2.2 as illustration of the intertemporal transformation of consumable units *ceteris paribus*, i.e. for a specific level of labor input and a specific physical capital stock. With a different level of human labor, or in the next period with a different stock of real capital, the intertemporal transformation will be characterized by a different production opportunity set.

[23] Admittedly, the exact numbers vary for different countries between 60%-40% to 70%-30%. Nevertheless, the expected stability of the functional distribution of income to human labor and physical capital will be a key insight for the framework; see Kaldor [1963], Valdés [1999, p. 5], or Barro and Sala-i Martin [2004, p. 12].

[24] von Hayek [1941] raised this question concerning the consistency of Fisher's [1930] view and Clark's [1899] theory of production, which underlies the equilibrium notion.

[25] This corresponds to the abstraction of fixed portfolio investments in housing, indicated in Sect. 2.2.2. The simplification can be justified on basis of the observation discussed in Chap. 1: in course of the Agricultural and Industrial Revolution land has been becoming increasingly less important for societies, so that mankind managed to escape the Malthusian doom.

Formal Modeling

Neoclassical growth theory enables a formal integration of these alternative production relations: following the classical contributions, the landmark article of Ramsey [1928] is widely considered as the starting point of the modern considerations. Harrod [1939] and Domar [1946] achieved an extension of the demand-determined equilibrium and its inherent instability in the spirit of Keynes [1936] into the long run. The mainstream of growth theory, however, is based on the contributions of Solow [1956] and Swan [1956], which allows to establish an extremely simple supply-side driven equilibrium model of the economy.[26] As explained, the basic building block of growth theory is an aggregate production function F of human labor $L(t)$ and physical capital $K(t)$:

$$Y(t) = F\left(K(t), L(t)\right), \tag{2.2.8}$$

where $Y(t)$ is the flow of consumable output produced at time t. Following Barro and Sala-i Martin [2004, pp. 26-28] neoclassical characteristics of this "consumption technology" materialize in three crucial properties: First, F exhibits constant returns to scale; i.e. mathematically, F is homogeneous of degree one:

$$F\left(\lambda \times K(t), \lambda \times L(t)\right) = \lambda \times Y(t). \tag{2.2.9}$$

Second, following the marginalists the aggregate production function is characterized by positive, but diminishing marginal products in either input factor:

$$\frac{\partial F}{\partial K} > 0 \qquad \frac{\partial^2 F}{\partial K^2} < 0 \qquad \frac{\partial F}{\partial L} > 0 \qquad \frac{\partial^2 F}{\partial L^2} < 0. \tag{2.2.10}$$

Ceteris paribus, i.e. holding the other factor constant, each additional unit of real capital or human labor delivers positive additions to macroeconomic output, but these additions decrease as the number of inputs increases.[27] And finally, the third characteristic of the neoclassical production function are the conditions of Inada [1963]:

$$\lim_{K \to 0} \frac{\partial F}{\partial K} = \lim_{L \to 0} \frac{\partial F}{\partial L} = \infty \quad \text{and} \quad \lim_{K \to \infty} \frac{\partial F}{\partial K} = \lim_{L \to \infty} \frac{\partial F}{\partial L} = 0. \tag{2.2.11}$$

Marginal productiveness of each input approaches infinity, as the input amount goes to zero, and approaches zero, when the amount goes to infinity.[28]

[26] For an extensive textbook treatment of growth theory see Jones [1997], Aghion and Howitt [1998], Valdés [1999], or Barro and Sala-i Martin [2004].

[27] Interestingly, Malthus [1815] was one of the first advocates of this so called *law of diminishing marginal productivity.*

[28] Frequently, *essentiality* of each input factor is considered as a fourth assumption of the neoclassical production function. However, this property is implied by the other three as shown in Barro and Sala-i Martin [2004, Sect. 1.5.1].

The usage of F as an explicit macroeconomic production function allows to *consistently* link the partial equilibriums in the markets for human labor, physical capital and consumable output into the broader idea of a long-run general equilibrium. Thus it exactly achieves the desired connection sketched in Sect. 2.1 and illustrated in Fig. 2.1: ignoring changes in the production technology itself, the only forces that can drive growth are changes in the labor input and the accumulation of physical capital. Due to the historical experience in demography the former influence has typically been restricted to the case of a population *growing* at a constant exogenous rate n,

$$L(t) = L_0 \exp(n \times t), \tag{2.2.12}$$

with L_0 as the initial population size. But this mathematical specification can also be used to address shrinking populations, i.e. $n < 0$. The increase of the stock of real capital is determined endogenously, addressing the discussed distinction of flows and stocks. Limiting this introduction to the simplest form of the growth model as initially given in Solow [1956], the net rate of increase is driven by the fraction of output saved and invested in real capital:

$$\frac{dK}{dt} = s \times Y(t). \tag{2.2.13}$$

There are several assumption underlying this relationship. Firstly, the economy is a closed one, so that all output is devoted to consumption or investments. As output corresponds also to income, the amount saved – i.e. not consumed – equals investments.[29] A second assumption in (2.2.13) is that this savings decision can be represented by a constant saving rate $s < 1$. One of the major tasks of Chapt. 3 will be to make this savings decision endogenous by including Sect. 2.2.1's fundamental trade off of current versus future consumption by the agents. Finally, depreciation is also ignored, or – following Solow [1956] – output is defined as net of depreciation. Neither does (2.2.13) allow for an adjustment cost effect as discussed in Sect. 2.2.4's presentation of the q-Theory. The replication framework will also incorporate these aspects.

[29]The assumption of a closed economy is also reflected in (2.2.12)'s limitation to the domestic labor force and the corresponding abstraction of international migration. Though globalization certainly enforces international interaction, the degree to which world capital markets are integrated, remains an open question. The effects of migration are expected to be limited as seen in Table 1.1. Based on the observations of Feldstein and Horioka [1980], Frankel [1991], and Taylor [1996], who document substantial correlation between national saving and national investment, and French and Poterba's [1991] home bias in equity investments, the strong assumption of a closed economy will thus be maintained for Pensionomics. This is also justified by the fact, that aging is a *global* phenomenon as Chap. 1 explained. The role of demography-driven international capital flows is addressed in Goyal [2002], Domeij and Floden [2003, 2005], Brooks [2003], Feroli [2003], Börsch-Supan et al. [2002], or Börsch-Supan et al. [2005].

Furthermore, the equality of total population size and labor force, which underlies (2.2.12), is another restrictive assumption to be given up, when old-age savings and thus the phenomenon of non-working retirees shall be addressed.

Basic Structure

It is convenient and insightful to rewrite the growth model in terms of per capita variables. This allows to directly address growth theory's typical concern of output and income per capita at a varying size of the population. As App. A.2 shows the fundamental equation of the Solow-Swan-Model in per capita is given by

$$\frac{dk(t)}{dt} = s \times f\left(k(t)\right) - n \times k(t). \qquad (2.2.14)$$

$k(t) \equiv K(t)/L(t)$ and $y(t) \equiv Y(t)/L(t)$ are per capita counterparts of the physical capital stock and macroeconomic output; $f\left(k(t)\right) \equiv F(k(t), 1)$ is the intensive form of the production function. The assumption of non-existing scale effects is evident in $y(t) = f\left(k(t)\right)$: production per person is only determined by the amount of real capital each person has access to. Figure 2.3 illustrates the workings of (2.2.14).[30]

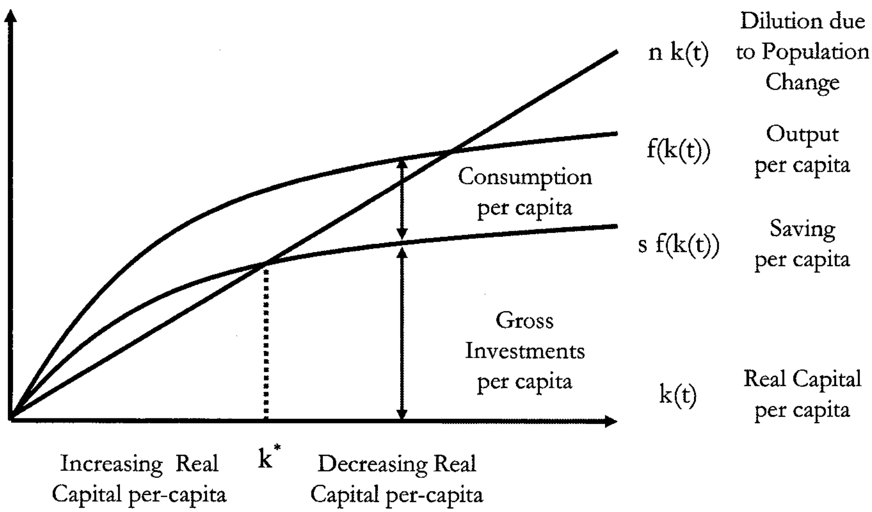

Fig. 2.3. Solow-Swan-Model

Remarks: The figure visualizes the neoclassical growth model of Solow [1956] and Swan [1956].

[30] An additional natural assumption here is that $f(0) = 0$.

The upper curve is the per capita production function $f(k(t))$, the lower one represents gross investments $s\,f(k(t))$, which are proportional to the production function. The shape of both is driven by the characteristic conditions of (2.2.11). The straight line from the origin represents the dilution of real capital per capita due to increases in the population. The change in $k(t)$ is given by the vertical distance between $s\,f(k(t))$ and $n\,k(t)$: right of the curves' intersection at k^* real capital per capita declines, since the population growth exceeds the investment-driven accumulation. Left of k^* it is the other way round, so that physical capital per capita increases. At k^* the two effects compensate each other. One has thus achieved a positive theory of the per capita capital stock and long-run per capita consumption.

These per capita relations are the most accessible representation of the economics, but the underlying effects are the same in absolute terms:[31] on the one hand, the labor input into the aggregate production function F changes from period to period, on the other hand, the input of physical capital is altered by depreciation and investments. The resulting output reflects both influences. Thereby, the Solow-Swan-Model can be used to explicitly incorporate the markets for human labor and physical capital. Based on Euler's adding-up theorem, the assumed homogeneity of F guarantees that the total output is exhausted by payments to the input factors, when these are paid their marginal products as implied by the neoclassic paradigm:

$$Y(t) = F(K(t), L(t)) = \frac{\partial F}{\partial K} \times K(t) + \frac{\partial F}{\partial L} \times L(t). \qquad (2.2.15)$$

Because under competitive markets, $\partial F/\partial L$ relates to the income from labor and $\partial F/\partial K$ is the compensation physical capital receives for its service in the production process, this allows for positive arguments on the wage and interest rate. However, the stylized facts of increasing stocks of physical capital per capita and stable interest rates levels cannot be rationalized within this simplest form of the framework. Diminishing returns, as implied by (2.2.11), would make it impossible to maintain per capita growth in the long run by just accumulating more physical capital per worker. While Chapt. 5 will explicitly address technological change as an additional element, the required extension can also be understood with the inclusion of a simple scaling effect into (2.2.8):

$$Y(t) = A(t) \times F(K(t), L(t)) = A(t)L(t)f(k(t)), \qquad (2.2.16)$$

The term $A(t)$ reflects the total factor productivity in the spirit of Solow [1957, Eq. (1a)] and can be understood as "shifts" in the production function. Positive shifts, i.e. increasing $A(t)$ over time, can now explain the observed facts that per capita output and income rose steadily, while interest rates remained relatively stable. For F the most frequently used functional form is

[31] While Chap. 5 will use the per capita form, the major part of the work in Chaps. 3 and 4 will be in absolute terms.

the celebrated Cobb-Douglas one, $Y(t) = A(t)\,K(t)^{\theta}\,L(t)^{1-\theta}$, with a constant $0 < \theta < 1$.[32]

Applications and Extensions

Classic contributions in the area of growth theory like Phelps [1961, 1965] focussed on the question of an optimal rate for growth. The very basic model discussed so far has been extended in several direction to improve its implications for the understanding of variations in per capita income across time and countries. In particular the role of research and development and of human capital formation in spirit of Becker [1960, 1962, 1993] has played a vital role – for instance in Becker et al. [1990] or in the Uzawa [1965]-Lucas [1988]-model. Ramsey [1928], Cass [1965] and Koopmans [1965] provided the link to the consumers' intertemporal optimization highlighted in Fig. 2.2.
While often ignored, uncertainty has increasingly been introduced into the discussion – above all by financial economists like Merton [1969*a*]. This provides also the link to the direct production-based approaches in the financial field. As mentioned earlier, the direct production models were of limited scope, as they considered a linear technology, which – though stochastic – focussed on constant returns to scale of real capital only. In contrast to this, contributions related to the idea of the neoclassical growth model, like Brock and Mirman [1972], Bourguignon [1974], Merton [1975], Brock [1979], Malliaris and Brock [1982], Brock [1982], or Chang and Malliaris [1987], clearly reflect the decreasing marginal productiveness of physical capital in their concave production functions.
Hence, the growth theoretic contributions meet easily the requirement of macroeconomic consistency and address the feedback of savings on production, but they have naturally hardly any focus on pension problems or portfolio diversification.

[32] Section 5.3.1 will fine-tune the understanding of $A(t)$ and clarify that the more justified modeling of technology in a Harrod [1939]-neutral form does *not* help in the pension problem. Therefore, the major part of the replication framework will be based on the simple interpretation as total factor productivity. Furthermore note, that economist Paul H. Douglas and mathematician Charles W. Cobb developed their famous production function explicitly to fit the empirical evidence for output, employment and the capital stock; see Barro and Sala-i Martin [2004, p. 29].

2.2.6 Public Finance

Finally, the last missing cornerstone for addressing the three dimensions set in Sect. 2.1 is a description of the unfunded pay-as-you-go mechanism as second pension system. The analysis in Sects. 2.2.1 to 2.2.5 has exclusively focussed on the capital market, where agents save and invest in order to live of the proceeds in the future. However, alongside this market based mechanism societies have developed other forms for reallocation resources across age and time – public pension systems being the most important ones.[33]

Overlapping Generations Model

Key to the understanding of an unfunded pension system is the concept of overlapping generations. While Samuelson [1958] and Diamond [1965] popularized it, Allais [1947] must be credited for proposing the "overlapping generations model" first. Like the life-cycle models for agents' intertemporal consumption smoothing, this framework focuses on the finiteness of human life and the consequences of this for economic development. In the model's most basic form life ends after two periods, and the model distinguishes thus two types of agents: those born in the previous period, representing the old generation and the young generation born in this period and forming the subsequent period's old. Typically, one also refers to them as cohorts, which connects the agents to their birth period. This birth period remains unchanged in course of the life-cycle, while an agent's classification into generations changes as time passes. However, at a particular point in time both concepts can be used interchangeably, as they unambiguously identify the group.
Based on Shell's [1971] first general perfect-foresight model with overlapping generations, early applications focussed on pure exchange economies and analyzed existence, uniqueness, stability and efficiency of its equilibrium as well as the special role of money.[34] Samuelson [1975*a*,*b*] connects the overlapping generations model to the growth literature. After Auerbach and Kotlikoff's [1987] application in the analysis of fiscal policies it has been frequently used in this area. In the finance field, as mentioned in Sect. 2.2.3, generational aspects gained momentum in the explanation of the Equity Premium Puzzle, but researches like Constantinides, Donaldson and Mehra [2002, 2005*a*,*b*] increasingly apply it to other aspects of financial economics. The heterogeneity of agents, which the overlapping generations model implies, plays a central role thereby. A continuous version of the overlapping generations model has also become popular, for instance in Blanchard [1985] or Weil [1989*c*].

[33]For a discussion of other reallocation mechanisms like health care see Lee [1994*a*,*b*].

[34]Important contributions include Gale [1973], Cass and Shell [1980], Balasko and Shell [1980, 1981*a*,*b*] and Weil [1989*a*].

Pay-as-you-go Pension System

With this knowledge of the working of the overlapping generations model, one can finally turn to the second main pillar of the pension issue, the unfunded pay-as-you-go system. As Sect. 1.2 illustrated, this form of old-age provision is currently covering a vast majority of the labor forces in most developed countries. Invented in Germany almost 120 years ago, the system has widely been adopted all over the world. In the past, the German PAYGO system has been extremely successful in providing retirement income and survived – unlike the stock market – two major wars, the Great Depression, and the re-unification. However, the upcoming demographic change will put great pressure on it as Chap. 1 emphasized.[35]

With the overlapping generations framework the PAYGO mechanism can easily be characterized: each period the payments paid to the currently retired, old cohort come from payments made by the working young cohort. Thus, each old generation is supported by funds from another generation and does *not* draw on previously accumulated funds. Though the contributions are made in a form similar to a payroll tax, PAYGO pension systems have a insurance-like character and are not pure governmental redistribution schemes: each agent's benefits are typically related to his contributions made over the individual life-cycle via a system of "earning points", "average indexed monthly earnings" or similar measures. By characterizing average net pension benefits relative to the average net wages prior to retirement in the so-called *replacement rates*, public pension plans often give the impression of offering a form of defined benefits. Recent downward adjustments in the replacement rates due to the demography-driven financing pressures have softened this view and highlight the risk associated in this pension mechanism. Initially, the institutional implementations of most important public pension systems – like the German or U.S. one – included also some part of funding, i.e. worker's contributions were partially invested in the capital market to co-finance later pension payments. However, since the deterioration of nominal wealth in course of the Great Depression and Hyperinflation and the German system's predominant investment into government bonds between the two world wars, the public pension systems are de-facto pure PAYGO ones. The formal switch caused a long debate and highlighted the Mackenroth [1952, 1957]-hypothesis, that all retirement expenditures have to be made from the *current* national income. Unlike for individual private savings, there is *no* intertemporal transfer of this *aggregate* income, so that the current one is the only source to finance social- and retirement-related payments. Assets of funded savings create the illusion of relative safety, because their valuation – reflecting the anticipation of the *entire* future income stream – is quantitatively of a different dimension then the periodic expenditures. While funded and unfunded pension systems make thus no difference in his setting, Mackenroth's analysis is problematic

[35] See Shiller [2003*a*] or Börsch-Supan and Wilke [2003] for some historical remarks on the German PAYGO system.

for Pensionomics as it is based on a constant national income. It thus neglects the possibility of economic growth. Furthermore, stochastic aspects and hence the portfolio perspective are also ignored.[36]
Historically, the post-war Baby Boom has fostered the functioning of the PAYGO mechanism and justified so the complete switch to unfunded system, because many young agents financed the pensions of relatively few retirees in the old generation. However, after this introductive effect the early scepticism – like Friedman [1962, pp. 182-189] – proved true, and many PAYGO systems do demand governmental subsidies in order to keep announced pension commitments, which many generations have trusted on, and keep contribution rates at supportable levels. Confronted with the prospective demographic change, some systems re-introduced the partial financing strategy in the 1980s – whereas the German system had to cope with consequences of re-unification. In particular the establishing of the Social Security Trust Fund in the U.S. substantially revived the academic debate as reflected for instance in the many contributions of Feldstein [1974, 1976, 1980, 1985, 1987]. Recently, the perspective to *fully* switch to a funded system – the "privatization of Social Security" in the U.S. – and the worsened shape of PAYGO systems due to the aggravated aging have produced a similar new interest in this field.[37]
However, the portfolio perspective on both systems as set in Sect. 2.1 is not addressed.

2.3 Summary and Survey of Approach

So far this chapter has review major theoretical contributions of finance and economics forming the methodical foundation of the replication framework, which will be developed in Chaps. 3 and 4. Before doing so, a survey of this unusual approach and a relation of it to the established theories is convenient.

Summary of Building Blocks

As Sect. 2.1 illustrated, this work's step into Pensionomics attempts to meet requirements with respect to three dimensions: macroeconomic consistency, the simultaneous addressing of funded and unfunded pension systems, and the portfolio consideration of risk and return. Table 2.1 summarizes the limitations in this regard of the theoretical approaches discussed in Sect. 2.2.
Though many contributions are related to the pension problem, none gives a satisfying answer concerning the pension problem, which most societies will face given the upcoming demographic change.

[36] See Börsch-Supan [2001] and Bubb and Zimmermann [2002, Chap. 6].
[37] See, for instance, the collections of Feldstein [1988], Feldstein and Siebert [2002], Feldstein and Liebman [2002], Campbell and Feldstein [2004] and many more contributions. Textbook treatments of public finance aspects and their relations to economics are given in Myles [1995], Cullis and Jones [1998], or Rosen [2002].

Table 2.1. Limitation of individual building blocks

	Macroeconomic Consistency	Pension Systems	Portfolio Perspective
Fisher Separation	Focus on investments in real capital only	No consideration of PAYGO system	Model without uncertainty
Portfolio Approach	Not addressed	PAYGO system hardly included	Funded savings' risk and return
Asset Pricing	Focus on investments in real capital only	PAYGO system not addressed	Funded savings' risk and return
q-Theory	Focus on investments in real capital only	Not addressed	Not addressed
Growth Theory	Predominant focus on growing populations	No explicit modeling	Not addressed
Public Finance	Implications of overlapping generations	Description of PAYGO system only	Not addressed

Remarks: The table summarizes the limitations of the theoretical approaches discussed in Sect. 2.2 with respect to the requirements for the pension problem.

Integrated Model

The general equilibrium framework developed in the following chapters will therefore try to integrate important aspects of them into a portfolio-based macro-perspective on the unfunded and funded pension arrangements. Core to the framework is a financial intertemporal portfolio consideration as established by Samuelson [1969] and Merton [1969*b*, 1971, 1973]. However, in combination with the overlapping generations framework to address age-based heterogeneity among the consumption-smoothing agents, the portfolio consideration will by be in discrete time, like in Fama [1970*b*] and Constantinides et al. [2002]. Though myopic utility assumptions make the problem rather simplistic compared to pure portfolio models, it still allows to capture the economics of diversification. The insights from asset pricing are used to derive relations for the available securities' risk-return characteristics, that are consistent to the underlying macroeconomic development with its demographic change and accumulation of physical capital. The importance of this link to macroeconomics is motivated by the insightful analysis of Diamond [2000] and the considerations in Bubb and Zimmermann [2002]. Thereby, the growth model and q-Theory's adjustment cost interpretation are integrated into a single capital adjustment technology closely following Abel [2003]. This growth-theoretic background allows to easily derive coherent formulations for the income from physical capital and human labor. However, the simultaneous inclusion of wages and the wage-driven PAYGO system into the portfolio consideration is the biggest challenge. The obstacle is the non-linearity problem shown in (2.2.4), which prevents the application of many standard tools and concepts from the financial theory. Following Merton [1983], it is overcome in two steps: a benchmark model and its replication.

Replication Methodology

First, Chap. 3 will *ignore* the moral hazard problem concerning human capital and *not* address the pay-as-yo-go pension system. Instead, it will be assumed that agents can freely trade the capitalized value of their future labor income, so that there is no fundamental difference between human and physical capital on the asset market. However, in the macroeconomic growth model, where a consumption technology drives the incomes paid to each of the input factors, real capital and human labor are clearly distinguished as rival inputs. Uncertainty in this production relation is the key driver for the portfolio consideration. Merton's [1983] analysis is extended by allowing for the accumulation of physical capital based on the adjustment technology from Abel [2003], and by including time-preference in the agents' intertemporal optimization. Based on this utility maximizing consumption-investment problem, Chap. 3 derives the agents' portfolios consisting of real capital securities and *hypothesized* human capital ones. These reflect the assumption of marketable human capital and allow for the standard techniques in the intertemporal portfolio consideration. The integration of the agents' optimal behavior into a general equilibrium notion allows to derive representations of maintainable consumption levels regardless of the population development and clearly reflecting the uncertainties in macroeconomic production.
Second, Chap. 4 uses the results of this utility-maximization as benchmark for a replication, in which the human capital is *not* tradable, but a *public sector* including the unfunded, pay-as-you-go pension system exists. This follows the interpretation of a second-best solution by Lipsey and Lancaster [1956]: the – counterintuitive – implication of the second best theory is, that when the efficiency condition for a Paretian optimum is violated in *one* market, additional violations allow to achieve more efficiency than the attempt to secure efficiency in as many markets as possible would do. The nontradability of human capital on the capital market is such a violation. The introduction of the public sector, is hence seen as a *second* violation of the first-best efficiency conditions, in order to off-set the initial market failure. To answer the pension problem, the PAYGO's contribution rate will be calibrated so, that agents of all age groups can achieve exactly those levels of consumption that they would have in the first-best solution with marketable human capital. In this

sense Chap. 3 serves as benchmark for the replication in Chap. 4.
The Pensionomics' portfolio optimization within the pension mechanisms is thus an indirect one: the PAYGO system is understood as a substitute for the tradability of human capital. Marketable future labor income would induce agents to optimize their wealth portfolios in order to participate in both factor incomes. The same risk diversification can also be realized with the multi-pillar pension approach. Since the first-best's allocations are the result of individual consumption smoothing over the stylized life-cycle in a production economy, the replication addresses the need for adequate, affordable and sustainable pension provisions. Hence, this step into Pensionomics meets the three requirements set in Sect. 2.1: the growth-theoretic, production-based derivation of factor incomes yields macroeconomic consistency and the replication framework is an indirect implementation of the portfolio perspective on PAYGO and funded savings.

Part II

The Model

3

Tradable Human Capital

Old age isn't so bad when you consider the alternative.

Kotlikoff and Burns [2004]: "The Coming Generational Storm", p. 1

In this chapter the first-best model with tradable human capital is developed. It serves as the benchmark for the replication by the PAYGO system. The model is based on an overlapping generations framework with three cohorts: the young, the middle-aged and the old. In essence, it will be shown, that with tradability of future wage income the young generation has a strong motive to partially diversify its exposure to labor income uncertainty and invest in real capital. Analogously, the middle-aged split their savings for old-age between physical and human capital. This leads to a participation of the old generation in labor income.

3.1 Population and Overview

Since the entire replication framework is developed in order to address the issue of changing demographic structures, the assumptions concerning the population are key for the model. The simple form of the overlapping generations model introduced in Sect. 2.2.6 must be extended to three cohorts for capturing the desired properties which characterize the demography. In addition to the description of the population this section will also give a synoptical outline of the framework's part with tradable human capital in order to ease understanding when each aspect is presented in more detail later on.

3.1.1 Overlapping Generations Framework

In the discrete-time economy that is considered agents live for three periods: they start as non-working young, form the economy's labor-force in the next period as middle-aged and retire when old. Consequently, there are three generations present at each period of time as illustrated in Fig. 3.1: In period t the cohort $j = t - 0$ born in period t represents the young generation. The cohort $j = t - 1$, born in period $t - 1$, is the working generation of middle-aged agents and cohort $j = t - 2$, born in $t - 2$, consists of retired old agents. In

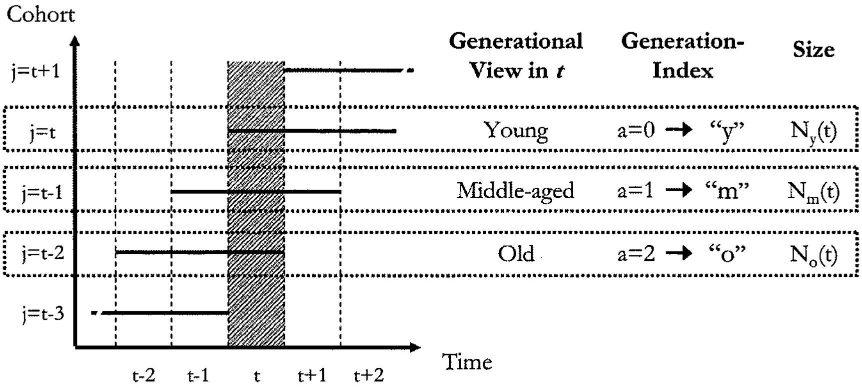

Fig. 3.1. Overlapping generations framework

Remarks: The figure illustrates the framework's assumption on population and demographics. In each period there are agents of three generations: a young cohort who has just been born into the economy and is not working, a middle-aged cohort representing period t's labor force and retired agents in the old cohort. In the course of the life-cycle the cohorts migrate through the different generations.

this sense j represent the birth period of a cohort that does not change over the life-cycle. As the naming of the overlapping *generations* model implies, it is useful to focus on the generational view: $a_j(t) = t - j$ characterizes the age of each cohort. Assuming that agents live for three period the values of to a_j is limited to $\{0, 1, 2\}$. Understandability can further be enhanced by using the indices "y" for $a_j = 0$ in case of young, "m" for $a_j = 1$ for middle-aged and "o" for $a_j = 2$ for old agents.

As indicated in Fig. 3.1 life expectancy is fixed and not uncertain in the model; i.e. all members of a cohort survive all three phases of their life-cycle. Given Chap. 1's understanding on trends in life expectancy this assumption is not too strong, but clearly focuses on declining fertility as the cause of aging. With life expectancy at birth exceeding 90 years, each period of time in the model represents approximately 30 years.[1] Furthermore, without premature mortality the size of each cohort does *not* vary over time. This implies that the generation sizes change over time according to

[1]Certainly most people start to work before they reach their thirties, yet the bulk of their working years is between the age of 30 and 60. Alternatively one could calibrate the setup to a life expectancy of 75 years with people entering the work force at the age of 25.

$$N_y(t) = \tilde{N}_y(t)\,, \tag{3.1.1}$$

$$N_m(t) = N_y(t-1)\,, \tag{3.1.2}$$

$$\text{and} \quad N_o(t) = N_m(t-1) = N_y(t-2). \tag{3.1.3}$$

While the size of the middle-aged, $N_m(t)$, and old generation, $N_o(t)$, is determined by the size of the younger generations in the previous period, the number of young agents is a realization of a positive random natural number: $\tilde{N}_y(t) \in \mathbb{N}$ This assumption makes decision on parenthood exogenous to the model and is based on the observation of Chap. 1 that birth rates are by far less predictable than mortality.[2]

As in Constantinides et al. [2002] *three* is the minimal number of generations – and thus life-cycle phases – capturing the desired features: a young generation faced with high uncertainty about its future prospects, a middle-aged one participating in the labor force and a retired one for whom uncertainty has mainly be resolved. Certainly, a higher number of overlapping generations – at the limit one per calendar year – would drastically reduce this major simplification of the model. Yet, three generations do already present a challenge in terms of analytical tractability as will be seen in this and the next chapter. On the other hand, if one limits the number of overlapping generations to only two, the desired participation of the old in the factor income of human labor is not achievable unless one allows for contracts with an unborn generation. Such a socially predetermined sharing rule would hardly follow the utility maximizing idea of the neoclassical paradigm and is thus not taken into consideration here.

All members within a given cohort are completely homogeneous; i.e. the identical agents have no idiosyncratic characteristics. This assumption corresponds to the remarks of Sect. 1.3, that the focus of this research on pensions is less that of intra-generational distributional effects than that of efficiency and inter-generational risk-sharing. With homogeneous agents in each cohort, one can methodologically reduce the set-up to a model of three representative agents, each symbolizing a different cohort. However, the possibility of a non-trivial demographic structure as implied by (3.1.1) to (3.1.3) makes this step inconvenient as the relative importance of the three cohort-representatives would have to be incorporated somehow differently. Therefore the layout of the framework will primarily be based on individual cohort members and only aggregated to a generation when required.

While heterogeneity within cohorts is downplayed, heterogeneity across generations is central to the model. The differences between generations stem from two facts: first, each generation is in a different phase of its life-cycle and thus one reasonably expects distinct decisions concerning savings and consumption, for instance. Second, with the introduction of a varying demographic structure and a certain process of accumulating real capital each generation's outlook on the future is different: young agent's cannot simply copy the de-

[2]However, (3.1.1) rules also out any insights from the New Home Economics.

cisions the middle took, when they were young, because these decisions have changed the state of nature. This consideration reflects the requirement for macroeconomic consistency emphasized in Chap. 2.

3.1.2 Outline of Model

With a discrete number of generations, the model is nested in a discrete time framework. Like in Fama's [1970*b*] analysis for multiperiod consumption-investment decisions in discrete time, a sequence of actions takes place in each period. This is visualized in Fig. 3.2.

At the beginning of a period all uncertainties are dissolved: the size of the young cohort – as introduced in this Sect. 3.1 – is observable and the realizations of the parameters in the macroeconomic production function, $\tilde{\theta}(t)$ and $\tilde{A}(t)$, completely determine the period's output. The description of the technology generating consumable output as well as the process for accumulating real capital are described in Sect. 3.2. This macroeconomic framework in real terms drives the real returns on the economy's capital market. Section 3.3 establishes the securities traded in this market: there are not only the usual riskless asset and securities tracking the value of physical capital, but also

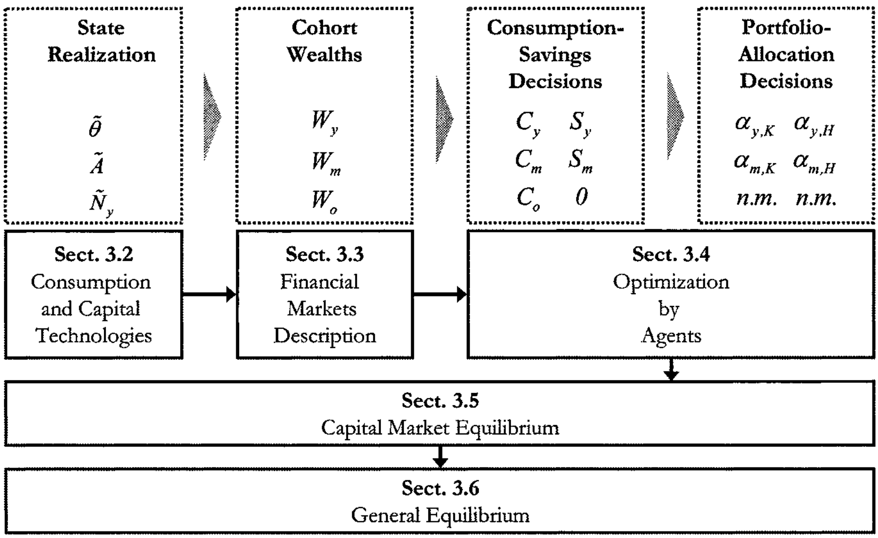

Fig. 3.2. Timing and outline

Remarks: The framework is a discrete time model: each period the dissolving of uncertainty, the determination of agents' wealth and the consumption and portfolio decisions happen subsequently without time passing. The figure also depicts how Chap. 3 is organized.

the hypothesized human capital securities representing claims on future labor income. The returns realized in the financial markets determine the wealth of each agent – and thus of each generation, which represents the budget for the consumption-savings decisions. The underlying intertemporal utility maximization is set up and solved in Sect. 3.4. This yields consumption policies and portfolio allocations for each cohort. In order to derive equilibrium in the capital market, the portfolio decisions are aggregated across generations in Sect. 3.5. The results are the aggregated amount of savings that is dedicated to human capital securities and the amount that is invested in physical capital. Section 3.6 establishes the corresponding equilibrium in the output market where investments in physical output and consumption must match the period's production of consumable units. This formally completes the model for the case of tradable human capital and Sect. 3.7 summarizes the main findings.

3.2 Consumption and Capital Technologies

In the closed economy considered here, there are two production technologies. The first is the *macroeconomic production* technology, in which physical capital and human labor are employed to produce consumption goods. These can either be consumed in the same period in which they are produced, or they can be used as input to the second, the *capital adjustment* technology. The capital adjustment combines consumption goods and existing physical capital in order to produce real capital units for the subsequent period of time.

3.2.1 Macroeconomic Production Function

The technology of producing consumption goods is of the frequently-used Cobb-Douglas type as introduced in Sect. 2.2.5:

$$\tilde{Y}(t) = \tilde{A}(t) \times K(t)^{\tilde{\theta}(t)} \times L(t)^{1-\tilde{\theta}(t)}, \tag{3.2.1}$$

where $Y(t)$ is aggregate output of consumption units in time period t. $\tilde{A}(t)$ is a positive random scaling factor reflecting the technical knowledge of the economy, independent from the stock of real capital, $K(t)$, and the aggregate amount of labor, $L(t)$, used in production. Uncertainty concerning the productivity of the input factors capital and labor is introduced by the random elasticity parameter $\tilde{\theta}(t)$, whose range is

$$0 \leq \tilde{\theta}(t) \leq 1. \tag{3.2.2}$$

Furthermore it is assumed that $\{\tilde{\theta}(t), \tilde{\theta}(t+1), \tilde{\theta}(t+2), \ldots\}$ are independent and identically distributed with an expected value of $\mathbb{E}_{t-i}[\tilde{\theta}(t)] \equiv \bar{\theta} \,\forall\, i = 1, 2, 3, \ldots$, where $\mathbb{E}_{t-i}[\circ]$ is the expectation operator conditional on knowing all information available in period $t - i$. This assumption will allow to address the *near* stability of the functional distribution of income to human labor and physical capital indicated in Sect. 2.2.5. Note that the consumption technology's specification in (3.2.1) is conform with the standard neoclassical setting as discussed in that section. Essentially, $\tilde{\theta}(t)$ presents the production elasticity of real capital and $1 - \tilde{\theta}(t)$ the production elasticity of labor. In contrast to standard growth theory, they are not constant but fluctuate randomly around fixed means. Yet, they always add up to one, maintaining the homothetic feature of the production function. $\tilde{A}(t)$ is also referred to as the total factor productivity.[3]

Output

Aggregate output of $Y(t)$ is measured in consumption units. These can either be consumed by the agents or saved and dedicated to the accumulation of physical capital. Thus, on the goods market non-consumption corresponds to savings being allocated to physical capital and entering the capital adjustment technology as investments. Let aggregate consumption be denoted by $C(t)$ and gross investments in physical capital by $I(t)$. Standard national income accounting reads

$$Y(t) = C(t) + I(t), \tag{3.2.3}$$

which must be fulfilled for equilibrium in the market of consumable output.

Firms

The economy's firms are in perfect competition and make all production decisions so as to maximize their market value. The firms' managers do not know the realizations of either $\tilde{A}(t)$ or $\tilde{\theta}(t)$ when they have got to make their investment decisions at time $t-1$, $I(t-1)$. However, when they determine the amount of labor to be employed, $A(t)$ and $\theta(t)$ have been revealed and the managers *do* know them. Consequently, the demand for labor at time t is determined by the managerial maximization problem with respect to profits, $\Pi(t)$: [4]

[3]Based on $\mathbb{E}_{t-i}[\tilde{\theta}(t)] = \bar{\theta}$, also the mean of labor's production elasticity is fixed at $\mathbb{E}_{t-i}[\tilde{\theta}(t) - 1] = \bar{\theta} - 1$. Furthermore, note that the assumption of $\tilde{A}(t) > 0$ excludes senseless negative total factor productivity. Finally, it is worth mentioning that this understanding of total factor productivity is based on Solow [1957], whereas Sect. 5.3 will address technical progress explicitly with a different implication for the corresponding variable.

[4]In addition to this, note that the managers' decision on the amount of investment is made on behalf of the shareholders. As mentioned in Sect. 3.2.1 this is done under uncertainty on $\tilde{A}(t)$ and $\tilde{\theta}(t)$. While this separation of ownership and control is an interesting relationship per se, it is not the focus of the model here. In accordance with the neoclassical paradigm it is thus assumed that managers act perfectly in the best interest of the shareholders.

$$\max_{L(t)} \Pi(t) \equiv p_Y(t) \times Y(t) - \omega(t) \times L(t). \qquad (3.2.4)$$

$\omega(t)$ refers to the wage rate and $p_Y(t)$ is the price of consumption goods. As nominal considerations play no role in this real world model, let the price of consumption goods be normalized to one, i.e. $p_Y(t) \equiv 1 \, \forall \, t$. Hence $\omega(t)$ is the *real* wage rate.

Equation (3.2.4) is the standard maximization of revenues minus costs, where the only costs to be addressed are labor. It is assumed that factor markets are competitive. Using (3.2.1) in (3.2.4) the first order condition of this optimization is[5]

$$A(t)K(t)^{\theta(t)}\,(1-\theta(t))\,L(t)^{-\theta(t)} - \omega(t) \overset{!}{=} 0$$

and by multiplication with $L(t)$

$$(1-\theta(t))Y(t) = \omega(t)L(t)\,.$$

Thus the aggregated demand for labor, $L^D(t)$, is

$$L^D(t) = \frac{(1-\theta(t))\,Y(t)}{\omega(t)}\,. \qquad (3.2.5)$$

and labor's share of output, i.e. the absolute factor income of labor, is given by

$$\omega(t)L^D(t) = (1-\theta(t))\,Y(t). \qquad (3.2.6)$$

Using this in the definition of profits, net revenues of all firms available for distribution to shareholders will be $\theta(t)Y(t)$, as normally in the productivity-based theory of income distribution.

3.2.2 Capital Adjustment Technology

The second technology of the economy governs the accumulation of real capital. Section 2.2.4 explained that with stochastic shocks in production the price of capital in terms of consumption units is *not* immutably be fixed at unity. The resulting deviation is incorporated into the framework by the capital adjustment technology: physical capital available for usage in the following period is produced by employing the unconsumed part of output – i.e. $I(t)$ – and the current real capital stock.[6]

[5] Naturally, the normalization $p_Y(t) \equiv 1$ has been applied as well. Otherwise one would obtain the notation of real wages here, $\omega(t)/p_Y(t)$.

[6] As Sect. 2.2.4 explained, the capital adjustment technology leads to a distinction between the valuation of unconsumed output and its valuation as input to the capital adjustment process. This will be addressed in the following.

Following Abel [2001, Eq. (5)] let the capital adjustment technology have a Cobb-Douglas specification, as well:

$$K(t+1) = B \times I(t)^{\vartheta} \times K(t)^{1-\vartheta} \tag{3.2.7}$$

where $B > 0$ and $0 \leq \vartheta \leq 1$ are scaling and elasticity parameters. Based on the Cobb-Douglas characteristics the capital adjustment technology exhibits the convexity of adjustments costs discussed in Sect. 2.2.4: additional investments have a positive but decreasing marginal effect on the physical capital stock. Section 5.2.5 will reflect further on this. But, though above formulation allows for an easy analytical tractability of the adjustment mechanism, it can only serve as an *approximation* to the more natural additive specification, as Abel [2001, p. 554] notes. Actually the capital adjustment technology should reflect a linear relationship between gross investment, depreciation and net investment, i.e. something like

$$K(t+1) = g(K(t), I(t)) + (1-\delta_K)K(t).$$

Here, investment net of adjustment costs, $g\left(K(t), I(t)\right)$, and depreciation, $\delta K(t)$, clearly add up to gross investment. While such a formulation is obviously more appealing as it reflects the common national income accounting conventions better, it has severe disadvantages for formal analytical work as will be seen later.[7] The specification of (3.2.7), however, circumvents these problems by its log-linearity and was hence used by Abel [2001] and Basu [1987].

Extreme Cases

Furthermore, the specification of (3.2.7) also nests the extreme cases of standard theory as discussed in Chap. 2. When $\vartheta = 1$ and $B = 1$ one obtains the *neoclassical "growth" model with complete depreciation* as referred to by Abel [2001, Footnote 5] i.e.[8]

$$K(t+1)|_{\vartheta=1} = I(t) \tag{3.2.8}$$

On the other hand, (3.2.7) also allows for a constant capital stock as in the Lucas [1978] fruit-tree model of asset pricing. For this the parameters must be set at $\vartheta = 0$ and $B = 1$ so that the capital adjustment technology degenerates to

$$K(t+1)|_{\vartheta=0} = K(t). \tag{3.2.9}$$

Thus the log-linear specification for the accumulation of real capital does not only have convenient analytical properties, as will be seen in the next subsection, but also serves as a nice generalization for standard models' implicit

[7] See Footnote 9.

[8] For convenience the setting of B to unity will be suppressed in the notation. Decisive is parameter ϑ.

assumptions on physical capital adjustment.
Section 5.2.5 will focus on the adjustment costs and growth aspects of this accumulation technology.

3.2.3 Price of Real Capital and Q-Ratio

Having established the capital adjustment technology, one is able to determine the price of real capital in terms of consumption units. This price $p_K(t)$ can be considered as the opportunity costs – measured in consumption units – of acquiring one unit of physical capital in period t to be carried into period $t+1$. In other words, it is the amount by which investments must be increased in order to increase next period's capital stock by a single unit, i.e. it is

$$p_K(t) \equiv \frac{\partial I(t)}{\partial K(t+1)}. \tag{3.2.10}$$

Due to (3.2.3), any increase of investments comes at the expense of current consumption. Applying the adjustment technology of (3.2.7) to this definition, the price of real capital is given by

$$\begin{aligned} p_K(t) &= \left(\frac{\partial K(t+1)}{\partial I(t)}\right)^{-1} = \left(B\vartheta \frac{K(t)^{1-\vartheta}}{I(t)^{1-\vartheta}}\right)^{-1} \\ &= \frac{1}{B\vartheta}\left(\frac{I(t)}{K(t)}\right)^{1-\vartheta}. \end{aligned} \tag{3.2.11}$$

According to (3.2.11), the price of a unit of physical capital depends on the amount of investment and on the installed real capital. Furthermore, as in Abel [2001] the value of the entire capital stock carried into the next period is $K(t+1) \times p_K(t)$, since it is acquired on today's output market. Using (3.2.7) and (3.2.11), this value – again measured in consumption units – can be determined as

$$\begin{aligned} K(t+1)p_K(t) &= \frac{1}{B\vartheta}\left(\frac{I(t)}{K(t)}\right)^{1-\vartheta} B\, I(t)^{\vartheta} K(t)^{1-\vartheta} \\ &= \frac{1}{\vartheta} I(t). \end{aligned} \tag{3.2.12}$$

Equations (3.2.11) and (3.2.12) reflect the analytical convenience of the capital adjustment technology's specification as Cobb-Douglas function in (3.2.7). The representation for the value of the real capital stock requires only a single additional parameter ϑ.[9]
Finally, note that the price of real capital as defined here reflects the phenomenon discussed as Tobin's [1969] q-ratio in Sect. 2.2.4. Notating it with $p_K(t)$, and not with an evenly intuitive "$q(t)$" is only for conformity with other price variables like p_Y or similar following in Sect. 3.3.

[9]Using instead of (3.2.7) the additive capital adjustment specification $K(t+1) = g(K(t), I(t)) + (1-\delta)K(t)$, would result in unappealing results for the price of real capital units and the value of the capital stock:

3.2.4 Real Capital's Rental and Return

While human labor is only employed in the consumption technology, physical capital enters both technologies as input factor. Hence, its service earns a rental in both technologies. Deriving an explicit formulation of these rentals instead of simply assuming an interest rate process reflects the economic foundation of the framework as demanded in Chap. 2.

Consumption Technology

It has already been established in Sect. 3.2.1 that in competitive factor markets net revenues of the firms are $\tilde{\theta}(t)Y(t)$. These profits can either be reinvested or distributed as dividends to shareholders. While the decision about this is up to the firms' annual general meetings, the value of the physical capital, i.e. the value of the firms, must not be affected by it. According to the theorem of Modigliani and Miller [1958] the dividend policy is irrelevant for the value in this purely neoclassical setting. This means that in the consumption technology each unit of real capital earns a rental of

$$\frac{\theta(t)Y(t)}{K(t)},$$

which is total factor income in the consumption technology payed to real capital divided by the units of real capital in place.

Capital Adjustment Technology

Following Abel [2001], the rental of physical capital in the capital adjustment technology is determined by its marginal product in this technology. The service of an additional unit increases the stock of physical capital in the following period. This increment must be valued at the corresponding price of real capital. Therefore, the rental is the marginal product of real capital in the capital adjustment technology – measured in terms of capital of the following period – multiplied by the price of real capital – measured relative to consumption units. In the capital adjustment technology, each unit of physical capital earns thus a rental of

$$\frac{\partial K(t+1)}{\partial K(t)} \times p_K(t) = B(1-\vartheta)I(t)^{\vartheta}K(t)^{-\vartheta}\frac{1}{B\vartheta}\left(\frac{I(t)}{K(t)}\right)^{1-\vartheta} = \frac{1-\vartheta}{\vartheta}\frac{I(t)}{K(t)},$$

where the result of (3.2.11) has been used.

$$p_K(t) \equiv \frac{\partial I(t)}{\partial K(t+1)} = \left(\frac{\partial K(t+1)}{\partial I(t)}\right)^{-1} = \left(\frac{\partial g(K(t),I(t))}{\partial I(t)}\right)^{-1}$$

$$p_K(t)K(t+1) = \frac{\partial I(t)}{\partial g(K(t),I(t))}g(K(t),I(t)) + (1-\delta)\frac{\partial I(t)}{\partial g(K(t),I(t))}K(t)$$

The analytical advantage of (3.2.11) and (3.2.12) is evident.

Total Rental and Return

Adding the results from above, one has the total rental of real capital in period t:

$$\frac{\theta(t)Y(t)}{K(t)} + \frac{1-\vartheta}{\vartheta}\frac{I(t)}{K(t)} = \frac{\vartheta\theta(t)Y(t) + (1-\vartheta)I(t)}{\vartheta K(t)} \tag{3.2.13}$$

As each unit of real capital in place in period $t+1$ has to be purchased in period t at the price of $p_K(t)$, the gross rate of return, $\tilde{R}_K(t+1)$, on holding a single unit of real capital from t to $t+1$ is

$$\begin{aligned}\tilde{R}_K(t+1) &= \frac{\frac{\vartheta\tilde{\theta}(t+1)\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}{\vartheta K(t+1)}}{p_K(t)} \\ &= \frac{\vartheta\tilde{\theta}(t+1)\tilde{Y}(t+1) + (1-\vartheta)\tilde{I}(t+1)}{\vartheta K(t+1)p_K(t)}.\end{aligned} \tag{3.2.14}$$

Because next period's realizations of $\tilde{A}(t+1)$ and $\tilde{\theta}(t+1)$ and thus $\tilde{Y}(t+1)$ as well as $\tilde{I}(t+1)$ are uncertain from period's t point of view, the return for holding a unit of real capital is risky.

3.2.5 Extreme Cases

It was shown in Sect. 3.2.2 that the log-linear specification of the capital adjustment technology nests the famous Lucas [1978] model as well as complete depreciation of real capital as extreme cases. This section examines the findings for prices, rentals and return on real capital for these extreme cases.

Lucas-Tree

The fruit-tree model of Lucas [1978] has the degenerate capital adjustment technology of $K(t+1) = K(t)$ where the parameters in (3.2.7) are set at $\vartheta = 0$ and $B = 1$. Taking the limit of (3.2.11) shows that a unit of real capital is invaluable in the Lucas-setting:

$$\lim_{\vartheta\to 0} p_K(t) = \lim_{\vartheta\to 0} \frac{1}{B\vartheta}\left(\frac{I(t)}{K(t)}\right)^{1-\vartheta} = \infty \tag{3.2.15}$$

The reason for this is straightforward: the fruit-tree of Lucas bears nonstorable fruits every period of time until infinity. The tree neither requires any investments nor depreciates. Therefore, the shadow price of a marginal unit of capital which would allow storing fruits is infinite in consumption terms.[10]

[10]Note bene: this is distinct of assessing a present value for the perpetual consumption stream from the tree, which is clearly not infinite. The price of physical capital rather captures the willingness to pay for the mere *existence* of a storage technology which is not present in the Lucas case. In the general case of (3.2.7) these two coincide as will be seen in Sect. 3.3.1: the marginal value of an additional unit of the storage device equals the present value of the future consumption payoffs realizable by it.

Consequently the rental and return of physical capital are useless concepts in this case. As described in Sect. 2.2.3, the Lucas-model is *not* a production-based approach but merely an endowment economy taking the endowment process as completely exogenous. When a marginal unit of real capital is infinitely valuable it will not be traded between different generations.
Therefore it does not yield additional insight to address the extreme case of Lucas [1978] any further.

Complete Depreciation

The other extreme of complete real capital depreciation – in other words, a model without accumulation of physical capital – was obtained when $\vartheta = 1$ and $B = 1$. Then the capital adjustment technology was found to be $K(t+1) = I(t)$ and the per-unit price of real capital can be determined from (3.2.11) as

$$p_K(t)|_{\vartheta=1} = \frac{1}{1 \times 1}\left(\frac{I(t)}{K(t)}\right)^{1-1} = 1. \tag{3.2.16}$$

This is the standard neoclassical result in *absence* of adjustment costs: with $\vartheta = 1$ and $B = 1$, units of physical capital can be de-installed and consumed without any losses in terms of consumption units. Each unit of such a fully consumable real capital stock has *exactly* the same value as a consumption unit, thus the relative price of real capital in terms of consumption is one. Equation (3.2.14) implies for the return on real capital

$$\begin{aligned} \tilde{R}_K(t+1)|_{\vartheta=1} &= \frac{1 \times \tilde{\theta}(t+1)\tilde{Y}(t+1) + (1-1)\tilde{I}(t+1)}{1 \times K(t+1)p_K(t)} \\ &= \frac{\tilde{\theta}(t+1)\tilde{Y}(t+1)}{K(t+1)} = \frac{\tilde{\theta}(t+1)\tilde{Y}(t+1)}{I(t)}, \end{aligned} \tag{3.2.17}$$

where in the last line $K(t+1)|_{\vartheta=1} = I(t)$ is used based on the formulation of the adjustment technology (3.2.8). Obviously, the case of complete depreciation is neither trivial nor useless. Given the fact that the time periods under consideration are very long due to the complexity of the overlapping generations framework, complete depreciation of real capital over a single period must not be considered unrealistic. Hence, each part of the model will also be rolled out for these extreme parameter values. Furthermore, this case without real capital accumulation is a simplified version of the model and will serve as the framework for an illustrative numerical example.

3.2.6 Illustrative Example Economy

Consider a simple economy without real capital accumulation, i.e. $\vartheta = 1$ and $B = 1$ in (3.2.7) and hence a capital stock equaling previous investments, $K(t+1) = I(t)$. The price of consumption goods in this \$-denominated economy shall be normalized to one, i.e. $p_Y(t) = \$1\,\forall\, t$. According to the reasoning of Sect. 3.2.5, the price of a unit of physical capital is then also fixed at \$1. Let the investments in the current period $t = 0$ be \$200 so that this constitutes the physical capital stock for $t = 1$. Assume that the size of the labor force employed in $t = 1$ is known to be 100 full-time working persons. Let there be two possible cases for next period's scaling and elasticity parameters each: $\tilde{\theta}(1) \in \{0.2, 0.4\}$ and $\tilde{A}(1) \in \{10, 20\}$.

According to (3.2.1), aggregate output of consumable or investable units in $t = 1$, $\tilde{Y}(1)$, is then characterized by four possible states, where the probabilities P_1, P_2, P_3, P_4 are of no interest here:[11]

$$\tilde{Y}(1) = \begin{cases} \$10 \times 200^{0.2} \times 100^{1-0.2} \approx \$1,148.70 & \text{with } P_1 \\ \$20 \times 200^{0.2} \times 100^{1-0.2} \approx \$2,297.40 & \text{with } P_2 \\ \$10 \times 200^{0.4} \times 100^{1-0.4} \approx \$1,319.51 & \text{with } P_3 \\ \$20 \times 200^{0.4} \times 100^{1-0.4} \approx \$2,639.02 & \text{with } P_4 \end{cases}$$

As derived in Sect. 3.2.5, the uncertain gross return on real capital is thus characterized by four possible realizations:

$$\tilde{R}_K(1) = \begin{cases} \dfrac{0.2 \times \$1,148.70}{\$200} \approx 115\% & \text{with } P_1 \\ \dfrac{0.2 \times \$2,297.40}{\$200} \approx 230\% & \text{with } P_2 \\ \dfrac{0.4 \times \$1,319.51}{\$200} \approx 264\% & \text{with } P_3 \\ \dfrac{0.4 \times \$2,639.02}{\$200} \approx 528\% & \text{with } P_4 \end{cases}$$

This represents the possible returns an investor could face when buying a unit of real capital in period 0.[12] While his savings-investment decision will be examined in Sect. 3.4, the next section extends the investor's investment possibilities beyond the now established participation in income from real capital.

---

[11]For completeness, let $0 \leq P_i \leq 1\,\forall i \in \{1,2,3,4\}$ and $P_1 + P_2 + P_3 + P_4 = 1$.

[12]As a single period refers to approximately 30 years the return of 230% would correspond to a real annual net return of roughly 3%.

3.3 Financial Markets Description

The capital market of the considered economy is characterized by three kinds of securities, each trading in a financial market. First, there are shares of stock representing ownership of real capital. Second, future labor earnings are tradable via an human capital security. And third, a riskless security offers a riskfree return.

3.3.1 Real Capital Security

While Sect. 3.2.1 gave some intuition on the value and return of real capital from an macroeconomic point of view, this section will elaborate on it from a financial perspective.

Ownership of Real Capital

The real capital security represents ownership of the economy's physical capital stock. It was established that real capital is entitled to a profit of $\theta(t)Y(t)$ in the consumption technology production and that the shareholders can decide on how much of this is distributed as dividend and how much will be reinvested. The reinvestment is crucial: even though physical capital does not depreciate in the general case, next period's capital stock must be acquired by the managers of the firms on behalf of the stockholders. Equation(3.2.12) identifies the value of this reinvestment as the amount of real capital stock carried into the following period times its price per unit:

$$K(t+1)p_K(t) = \frac{1}{\vartheta}I(t).$$

Standard accounting on the firms' level implies that aggregate dividends paid, $D(t)$, net of reinvestments, $K(t+1)p_K(t)$, are given by

$$D(t) = \theta(t)Y(t) - K(t+1)p_K(t)\,. \tag{3.3.1}$$

Combining this with previous equation allows to identify $(1/\vartheta)I(t)$ as the *ex-dividend* market value of the firms' shares measured in consumption units.

$$\frac{1}{\vartheta}I(t) = K(t+1)p_K(t) = \theta(t)Y(t) - D(t)\,. \tag{3.3.2}$$

After paying dividends, ownership of the firms is worth the amount of investment realized corrected for the adjustment costs.

To understand this important result one must keep in mind that, even with adjustment costs, owners could decide to completely liquidate the firms and consume the proceeds. Of course, the re-transformation from physical capital into consumption units would be at a "discount" compared to the valuation

of installed capital reflecting the convex adjustment costs, but it is possible. Hence, the real capital security is only a medium to postpone consumption to the next period, where the decision has to be made again. Continued existence of the firms is then the result of sequential decisions of non-liquidation as owners find it beneficial to maintain the firms. In other words, the neoclassical assumptions allow to consider incremental changes in the firms' physical capital stock with the corresponding financing completely equivalent to a liquidation of the firms and re-foundation of them at the desired level of real capital stock. Under the Modigliani and Miller [1958] theorem the financing of the physical capital stock carried into the next period is irrelevant – yet it must be made at expense of current consumption.
To ease understanding of this derivation of the real capital security's price, it is is convenient to turn to finance's standard concept of prices as risk adjusted net present values of future payoffs. The payoffs for buying physical capital in t are the $t+1$-rentals in the production and capital adjustment technology. As physical capital could be completely de-installed and sold at the real price of consumption at this point of time, there are *no* further future claims from the investment in t. For the valuation of real capital it is not important whether the payoff is in form of distributed dividends or capitalized profits. Formalizing this, (3.2.14) could also be considered as a present value statement for real capital investment:

$$K(t+1)p_K(t) = \mathbb{E}_t\left[\frac{\tilde{\theta}(t+1)\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}{\tilde{R}_K(t+1)}\right]. \qquad (3.3.3)$$

The market value of the real capital securities, $K(t+1)p_K(t)$, is the expected present value of its future and thus uncertain rentals, $\tilde{\theta}(t+1)\tilde{Y}(t+1)+\frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)$ as derived in Sect. 3.2.4, discounted at the corresponding risk-adjusted rate of $\tilde{R}_K(t+1)$.
Finally, there is another important insight in the result of (3.2.12): apart from the degenerate case $\vartheta = 1$, for which the physical stock is worth exactly the investments realized, the valuation of it is higher than the mere investments realized. In other words, the firms are worth *more* than the investments made by them, as for $0 < \vartheta < 1$

$$K(t+1)p_K(t) > I(t)\,.$$

This possibly surprising result is a direct consequence of the convex capital adjustment technology: physical capital, once installed, has a value *per se* exceeding the amount of consumption that is given up – i.e. saved – and converted into real capital. Via the adjustment technology installed physical capital allows for consumption transformation not only between two subsequent periods, but through the entire time, even though on a diminishing base. In other words, real capital is not a trade-off of today versus tomorrow, but of today versus tomorrow and the day after tomorrow and so on. The mere possibility of this consumption transfer beyond tomorrow is valuable and hence earning a rental. This justifies the rental in the capital adjustment technology as defined in (3.2.13).

National Income Accounting

The last observation has a direct consequence for national income accounting. The capital adjustment technology adds value to the standard macroeconomic output, $Y(t)$. Therefore, the economy's entire generated value is rather measured by the gross domestic product, $Y^{GDP}(t)$, than by aggregate output alone. As in Abel [2001], the gross domestic product equals the market value of output dedicated to consumption, $C(t)p_Y(t)$, plus the market value of real capital carried into the next period, $K(t+1)p_K(t)$. Remembering the normalization of $p_Y(t) = 1$ and applying (3.2.12) the gross domestic product is

$$\begin{aligned} Y^{GDP}(t) &= C(t)p_Y(t) + K(t+1)p_K(t) \\ &= C(t) + \frac{1}{\vartheta}I(t) = C(t) + I(t) + \frac{1-\vartheta}{\vartheta}I(t) \\ &= Y(t) + \frac{1-\vartheta}{\vartheta}I(t), \end{aligned} \tag{3.3.4}$$

where the last line follows from (3.2.3). Equation (3.3.4) shows that $Y^{GDP}(t)$ exceeds $Y(t)$, if $\vartheta < 1$. Again this is due to the adjustment costs introduced by the capital adjustment technology. If there were no such costs, $I(t)$ would represent not only the input to the capital adjustment technology but also the value of its output. However, with the convexity of (3.2.7), gross domestic product exceeds output because $I(t)$ now measures only the *input* to the capital adjustment technology and *not* the value of this technology's output, namely installed physical capital. The conversion of saved consumption units, $I(t)$, into capital in place, $K(t+1)$, corresponds to the value of real capital per se.

3.3.2 Human Capital Security

The second security traded on the financial markets is of hypothetical nature. It is designed such that it perfectly tracks the value of human capital. This means that a worker does actually not receive labor income, but the payoff of the human capital securities he holds.

Hypothetical Nature

The assumption that these securities exist is crucial for the entire framework in terms of solution approach and results. In essence, the existence of a liquid market for the value of human capital enables the young to issue securities as claims on their future labor income in order to consume out of the proceeds. This also allows to diversify into other investments as will be seen later. The hypothetical securities separate the service that human labor delivers in the production of consumable output from the suppliers – the workers. If a young agent completely sells the capitalized value of his labor income, he still contributes the service of human labor to the production process, (3.2.1), but does not earn a factor income in form of a wage. Instead, the owner of the corresponding security receives a dividend for his investment corresponding to the wage. From a macroeconomic point of view, one can interpret the production of consumable output as a combination of services of two forms of capital: human one and physical one. The costs associated for doing so are the rentals for each factor: wages and profits. From a methodical perspective the human capital securities allow to fully employ the standard toolkit of finance and to circumvent the labor-caused problem of non-additivity in the intertemporal wealth budgets as discussed in Sect. 2.2.2.

Labor Market

In order to derive the payoff structure of these securities, one has to establish equilibrium on the labor market first. In Sect. 3.2.1 it was shown that aggregated demand for labor, $L^D(t)$, is given by

$$L^D(t) = \frac{(1-\theta(t))Y(t)}{\omega(t)}.$$

Supply of labor is determined by the working generation. Based on the assumptions of the overlapping generations framework in Sect. 3.1.1, only the middle-aged cohort works. Ignoring leisure completely, this generation dedicates its *entire* time to working. This might be considered unrealistic, but will allow major simplifications concerning utility, human capital and the replication in Chap. 4.[13] Without a wage rate dependent decision on the amount of time a cohort's members work, aggregate supply of labor equals the size of the working generation multiplied by the time worked per person. Without loss of generality, let this working time per person be normalized to one. In other words, it is measured in life-time working units. Labor supply, $L^S(t)$, is thus given by

$$L^D(t) = N_m(t). \tag{3.3.5}$$

Wage inelastic supply implies that equilibrium in the labor market is achieved at

$$L(t) = L^D(t) = L^S(t) = N_m(t)\,, \tag{3.3.6}$$

where $L(t)$ can be specified as the equilibrium amount of human labor. It has already been shown in (3.2.6) that labor's share of output is

[13]Nevertheless, interesting insights into the design of the PAYGO system can be gained by addressing the possibility of leisure. This will be done in Sect. 5.4.2.

$$\frac{\omega(t)L^D(t)}{Y(t)} = 1 - \theta(t).$$

Using this and (3.3.6), a period's equilibrium wage rate, $\omega(t)$, can be determined as

$$\omega(t) = (1 - \theta(t))\,\frac{Y(t)}{N_m(t)}. \tag{3.3.7}$$

Return on Human Capital

As human labor is only employed in the consumption technology but not in the capital adjustment technology, (3.3.7) shall be the payoff of each security tracking human capital. This means that a human capital security bought in period t represents a claim on

$$\tilde{\omega}(t+1) = \left(1 - \tilde{\theta}(t+1)\right)\frac{\tilde{Y}(t+1)}{\tilde{N}_m(t+1)}. \tag{3.3.8}$$

It must be reemphasized that the hypothetical human capital security is a surrogate for income from human labor. Even before an agent works, he can trade his human capital. In financial terms, it is the present value of future income from labor. Therefore, the human capital security represents a claim *on* $\tilde{\omega}(t+1)$ rather than *of* an $\tilde{\omega}(t+1)$-amount.
Let the period t-price of one share of this security be denoted with $p_H(t)$. Note that the size of the young cohort in one period completely determines the size of the labor force in the subsequent period which is thus not uncertain. In analogy to (3.2.14), the gross rate of return, $\tilde{R}_H(t+1)$, on holding a single security of the human capital kind from t to $t+1$ is

$$\begin{aligned}\tilde{R}_H(t+1) &= \frac{(1-\tilde{\theta}(t+1))\frac{\tilde{Y}(t+1)}{\tilde{N}_m(t+1)}}{p_H(t)}\\ &= \frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{N_m(t+1)p_H(t)}.\end{aligned} \tag{3.3.9}$$

Finally, note that investment in human capital in this model does *not* refer to individual or governmental spending on higher eduction or practical training! The simplicity of the labor market and the inelasticity of labor supply towards the wage rate excludes any enhancement of individual wage expectations by such *formation* of human capital. In this framework, investment in human capital refers always to the perspective of an investor allocating wealth into securities of the human capital sort.

Parenthood

Before introducing the third security, some considerations on the creation of human capital are mandatory. In analogy to the capital adjustment technology of Sect. 3.2.2, one could have tried to formulate a "technology" generating next period's human labor like for instance

$$L^S(t+1) = h(L((t), S^P(t)) + (1 - \delta_H)L(t),$$

where $h(L((t), S^P(t))$ is a kind of investment in parenthood net of some adjustment costs based on consumption foregone by the parents, $S^P(t)$, and $\delta_H L(t)$ is the "depreciation" of existing labor. Such a specification is clearly achievable under two conditions: First, each cohort must work more than a single period, otherwise existing labor decays completely, i.e. $\delta_H = 1$. Second, the model must be able to account for parental non-consumption that is invested in rising their children. Neither of these is met given the layout of the model here.[14]
The hypothesized labor adjustment technology is rather of a form similar to the Lucas [1978] case of the capital adjustment technology:

$$L^S(t+1) = N_m(t+1) = N_y(t)$$

However, as the process generating the cohort sizes is assumed to be exogenous, such a labor adjustment technology implies that parenthood does not have any opportunity costs in terms of reduced consumption. Unfortunately, this leads to a philosophically very unappealing result: parenthood does *not* imply the creation of any value per se – the fruitful it might be socially, economically and above all emotionally! A new cohort is simply added to the economy at no opportunity costs and hence without any value in addition to its aggregate human capital.[15]
Consequently, the definition of gross domestic product must not be supplemented by an additional component covering the creation of the young in the next period. Instead, proper accounting for the human capital generated requires to extend the strict framework of national income accounting to the concept of wealth as will be seen later.

3.3.3 Riskless Security

The last security in the capital market is the standard riskless asset. Holding it from t to $t+1$ offers a riskfree gross return of $R(t)$. Because the future payoff is not risky it is completely known when such a security is purchased.

[14] In addition, it seems unlikely that the decision to have children at the economic expense of reduced personal consumption is only and fully a rational one. Unfortunately, however, exactly this sort of consumption-egoism has contributed to the declining birth rates described in Chap. 1 as reasoned by the New Home Economics.

[15] To be very clear on this: Of course there is value generated by the emergence of a new generation, namely its human capital, i.e. the present value of its future labor income. However, in contrast to the explanations in Sect. 3.3.1, there is no value created *in addition* to this as seen for the capital adjustment technology when the parameter ϑ is smaller than one.

Following Merton [1983], the timing convention implies then the usage of t. As the model is in real terms, one could have the objection that a riskless real return does not exist, in particular not for the time to maturity which each time step represents here.[16] Still, the model manages this aspect by the fact that the riskless security is in zero net supply. This means it is *not* assumed that there exists a device generating the riskfree return from one period to another. The asset is merely a concept which allows financial market participants to trade a riskless payoff. Its introduction to the model enables the usage of standard tools from well established capital market theory. In particular, it allows for important insights to the risk considerations in saving decisions.
There is no ad-hoc specification for $R(t)$. Instead, the riskless return will be developed endogenously within the model in a sort of normative term structure model.

3.3.4 Capital Market and Summary

Having characterized all the securities traded in the economy's financial markets, one can aggregate them to the capital market.

Supply of Securities

As mentioned in the last section, the riskless security is in zero net supply. The number of outstanding human capital securities corresponds to the aggregate human capital of the economy. With the assumptions of the overlapping generations framework and the details on the labor market of Sect. 3.3.2, the entire young cohort, who is next period's labor force, determines the aggregated human capital. This means there are $N_y(t) = N_m(t+1)$ securities of the human capital kind.[17] Finally, let each unit of real capital be represented by a single security of the real capital kind, so that there are $K(t+1)$ of them in period t. Note that this is consistent with Sect. 3.3.1's notion of managers purchasing next period's physical capital: in order to do so they have to issue $K(t+1)$ shares of the real capital security.[18]

---

[16]However, recently inflation indexed bonds have come into vogue by international treasuries. While France or United States have already issued inflation indexed bonds, Germany is preparing the legal basis for doing so. With maturities of around 30 years these issuances present an obvious candidate for the riskless security in real terms. For further details on the importance of real versus nominal riskless returns see Campbell and Viceira [2001*a*].

[17]Note that this refers to the gross number of human capital securities. As Sect. 3.4 will show, the young agents sell their human capital securities only *partially* in the market and keeps some in their portfolios of savings.

[18]Of course, the reference to a single security per unit of real or human capital presents a normalization not limiting the model's generality. One could also introduce a certain number of securities outstanding per claim, but this would only introduce unnecessary complexity.

Return of Market

With this supply of securities, the return on the entire capital market can be derived easily: as usual in finance, the return on the capital market is a weighted sum of returns on the real, human and riskless securities. The weights correspond to the fraction each kind of security contributes to the value of the entire capital market. This "total market capitalization" in real terms is easily derived as $K(t+1)p_K(t) + N_m(t+1)p_H(t) + 0$. Hence, using the above assumptions on the securities' supply and (3.2.14) as well as (3.3.9) the gross rate of return on the capital market, $\tilde{R}_M(t+1)$, is

$$\begin{aligned}
\tilde{R}_M(t+1) = & \frac{K(t+1)p_K(t)}{K(t+1)p_K(t) + N_m(t+1)p_H(t) + 0} \times \tilde{R}_K(t+1) \\
& + \frac{N_m(t+1)p_H(t)}{K(t+1)p_K(t) + N_m(t+1)p_H(t) + 0} \times \tilde{R}_H(t+1) \\
& + \frac{0}{K(t+1)p_K(t) + N_m(t+1)p_H(t) + 0} \times R(t) \\
= & \frac{\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta} I(t+1)}{K(t+1)p_K(t) + N_y(t)p_H(t)}. \qquad (3.3.10)
\end{aligned}$$

One could also derive this result based on the observation that the aggregated payoff of the capital market is the entire generated value of the economy, i.e. the gross domestic product. Then the market return can be calculated as future payoff divided by current price. Hence $\tilde{R}_M(t+1)$ is also obtained by using above notion of the total market capitalization and applying the findings for $Y^{GDP}(t)$ from (3.3.4) shifted by one period into the future:

$$\tilde{R}_M(t+1) = \frac{\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta} I(t+1)}{K(t+1)p_K(t) + N_y(t)p_H(t)}.$$

Finally, Table 3.1 summarizes all the definitions and findings for the three securities and the capital market.

3.3.5 Full Depreciation

Based on the summary of Table 3.1, it is easy to derive the limiting case of fully depreciating real capital. As Sect. 3.2.5 showed this case is characterized by $\vartheta = 1$ and $B = 1$ in the capital adjustment technology. It has already been shown that then $p_K(t)|_{\vartheta=1} = 1$ and

$$\tilde{R}_K(t+1)|_{\vartheta=1} = \frac{\tilde{\theta}(t+1)\tilde{Y}(t+1)}{I(t)},$$

which is consistent with using $\vartheta = 1$ in the formulation for the real capital return in Table 3.1 and the limit for $K(t+1)$ – equaling $I(t)$ – from (3.2.8). The

Table 3.1. Financial markets summary

	Real Capital Security		Human Capital Security	
Supply	$K(t+1)= BI(t)^{\vartheta}K(t)^{1-\vartheta}$	(3.2.7)	$N_m(t+1) = N_y(t)$	Sect. 3.3.4
Price	$p_K(t) = \frac{1}{B\vartheta}\left(\frac{I(t)}{K(t)}\right)^{1-\vartheta}$	(3.2.11)	$p_H(t)$	Sect. 3.3.2
Payoff	$\frac{\tilde{\theta}(t+1)\tilde{Y}(t+1)+\frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}{K(t+1)}$	Sect. 3.2.4	$\tilde{\omega}(t+1) = \frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{\tilde{N}_m(t+1)}$	(3.3.8)
Return	$\tilde{R}_K(t+1) = \frac{\tilde{\theta}(t+1)\tilde{Y}(t+1)+\frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}{K(t+1)p_K(t)}$	(3.2.14)	$\tilde{R}_H(t+1) = \frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{N_m(t+1)p_H(t)}$	(3.3.9)
	Riskless Security		**Capital Market**	
Supply	0	Sect. 3.3.3	*n.m.*	Sect. 3.3.4
Price	1	Sect. 3.3.3	$K(t+1)p_K(t)+ N_y(t)p_H(t)$	Sect. 3.3.4
Payoff	$R(t)$	Sect. 3.3.3	$\tilde{Y}^{GDP}(t+1) = \tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta}I(t+1)$	Sect. 3.3.4
Return	$R(t) = \frac{R(t)}{1}$	Sect. 3.3.3	$\tilde{R}_M(t+1) = \frac{\tilde{Y}(t+1)+\frac{1-\vartheta}{\vartheta}I(t+1)}{K(t+1)p_K(t)+N_y(t)p_H(t)}$	(3.3.10)

Remarks: The table summarizes all findings for the three securities and the capital market. Note, how the $t+1$-payoff dividend by t-price gives the return from t to $t+1$ for each item.

return on the human capital security in absence of real capital accumulation is

$$\tilde{R}_H(t+1)|_{\vartheta=1} = \frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{N_m(t+1)p_H(t)}\,. \tag{3.3.11}$$

i.e. there is no difference compared to (3.3.9). Lastly, the return on the capital market with full depreciation can easily be specified as

$$\tilde{R}_M(t+1)|_{\vartheta=1} = \frac{\tilde{Y}(t+1)}{I(t)+N_y(t)p_H(t)}\,. \tag{3.3.12}$$

Furthermore, note that with $\vartheta = 1$, the standard neoclassical model reemerges in (3.3.4) with $Y^{GDP}(t) = Y(t)$. As in the last section, the return on the capital market could also have been derived by dividing the capital markets payoff – in this case $Y(t)$ – by the total market capitalization, which is now given by $I(t)+N_y(t)p_H(t)$ due to the mentioned limits for $K(t+1)$ and $p_K(t)$. Of course, this yields the same result as above.

3.3.6 Illustrative Example Economy

Continuing the illustrative example, one can calculate further aspects of this simple economy with full real capital depreciation. In Sect. 3.2.6 it was shown that next period's output is characterized by four possible states each corresponding to a specific return on real capital.
Similarly, there are also four states for the return on the human capital. Using the results on $\tilde{Y}(1)$ and assuming that the human capital security is currently trading at $p_H(0) = \$4.06$, these states can be derived from the results in the last section. According to (3.3.11) the possible realizations of the return on the human capital security are

$$\tilde{R}_H(1) = \begin{cases} \dfrac{(1-0.2)\times \$1,148.70}{\$4.06\times 100} \approx 226\% & \text{with } P_1 \\ \dfrac{(1-0.2)\times \$2,297.40}{\$4.06\times 100} \approx 453\% & \text{with } P_2 \\ \dfrac{(1-0.4)\times \$1,319.51}{\$4.06\times 100} \approx 195\% & \text{with } P_3 \\ \dfrac{(1-0.4)\times \$2,639.02}{\$4.06\times 100} \approx 390\% & \text{with } P_4 \end{cases}$$

Having specified both the return on the real and on the human capital security, one can proceed and calculate the possible capital market returns. Again referring to the no-accumulation results in Sect. 3.3.5, the assumed investment of \$100 and the cohort size of 200 yields according to (3.3.12)

$$\tilde{R}_M(1) = \begin{cases} \dfrac{\$1,148.70}{\$100+\$4.06\times 200} \approx 190\% & \text{with } P_1 \\ \dfrac{\$2,297.40}{\$100+\$4.06\times 200} \approx 379\% & \text{with } P_2 \\ \dfrac{\$1,319.51}{\$100+\$4.06\times 200} \approx 218\% & \text{with } P_3 \\ \dfrac{\$2,639.02}{\$100+\$4.06\times 200} \approx 435\% & \text{with } P_4 \end{cases}$$

Observe that this could also be derived by the market weighted sum of the individual securities' returns. For instance, for the first state one would have

$$\begin{aligned} \tilde{R}_M(1)|_{State\ 1} &= \frac{\$100}{\$100+\$4.06\times 200}\times 115\% + \frac{\$4.06\times 200}{\$100+\$4.06\times 200}\times 226\% \\ &\approx 38\% + 152\% = 190\%. \end{aligned}$$

This illustrates the consistency of the model on a numerical basis.

3.4 Optimization by Agents

Aligned with the neoclassical paradigm all agents of the economy maximize expected utility from consumption. However, due to the life-cycle effects introduced by the overlapping generations framework this optimization differs according to an agent's age resulting in heterogeneity across the cohorts. As indicated in Fig. 3.2 each agent must make two decisions: he has got to decide how to split his available wealth into current consumption and savings and he has got to allocate this as investments to the different securities available. The former is the consumption-savings decision, the latter the portfolio selection problem.

3.4.1 Utility Assumptions

The utility rationale will be introduced in two steps: First, it is shown how the fact that life is finite alters the utility considerations throughout the life-cycle on a very general level. Second, the notation is specified to the easily trackable form of logarithmic utility in the three generations model. The corresponding reduction in complexity is a justification for using the simple log-utility case.

Utility over Life-Cycle

Assuming that all agents have a time separable lifetime utility function with respect to consumption when they are born, let this utility of an agent who is in period t of age a be denoted $U_a(t)$.[19] Extending the description of Sect. 2.2.1, it is given by

$$U_a(t) = \mathbb{E}_t\left[\sum_{i=a}^{T} \beta^{i-a} \times u_i\left(c_i(t+i-a)\right)\right] \text{ for } a = 0, 1, \ldots, T-1, T, \quad (3.4.1)$$

where the framework must not be limited to the three generations model presented so far, but could be of more general kind with $T+1$ representing the constant and known life expectancy.[20] In (3.4.1), *period* utility of consumption, $u_a\left(c_a(t)\right)$, is an age-dependent function transforming consumption of an agent aged a in period t, $c_a(t)$, into units of utils. These are corrected for the time in which they occur by the corresponding power of the subjective discount factor β. $\sum$ is the common summation operator reflecting time separability of $U_a(t)$.

Observe how the lifetime utility function of (3.4.1) materializes for agents of different ages:

---

[19]The assumption that only consumption gives utility eliminates any bequest motives. Moreover, certainty of the life span also prohibits legacies due to sudden death.

[20]Each agent is born at age 0 and dies after period T giving him $T+1$ periods of life.

$$
\begin{aligned}
U_0(t) &= \mathbb{E}_t\left[\sum_{i=0}^{T}\beta^{i-0}u_i\left(c_i(t+i-0)\right)\right] \\
&= u_0\left(c_0(t)\right) + \mathbb{E}_t\left[\begin{array}{l}\beta u_1\left(c_1(t+1)\right) + \beta^2 u_2\left(c_2(t+2)\right) + \ldots \\ +\beta^{T-1}u_{T-1}\left(c_{T-1}(t+T-1)\right) \\ +\beta^T u_T\left(c_T(t+T)\right)\end{array}\right] \\
U_1(t) &= \mathbb{E}_t\left[\sum_{i=1}^{T}\beta^{i-1}u\left(c_i(t+i-1)\right)\right] \\
&= u_1\left(c_1(t)\right) + \mathbb{E}_t\left[\begin{array}{l}\beta u_2\left(c_2(t+1)\right) + \beta^2 u_3\left(c_3(t+2)\right) + \ldots \\ +\beta^{T-2}u_{T-1}\left(c_{T-1}(t+T-1)\right) \\ +\beta^{T-1} u_T\left(c_T(t+T)\right)\end{array}\right] \\
&\vdots \\
U_{T-1}(t) &= \mathbb{E}_t\left[\sum_{i=T-1}^{T}\beta^{i-T+1}u_i(c_i(t+i-T+1))\right] \\
&= u_{T-1}\left(c_{T-1}(t)\right) + \mathbb{E}_t\left[\beta u_T(c_T(t+1))\right] \\
U_T(t) &= \mathbb{E}_t\left[\sum_{i=T}\beta^{i-T}u_i(c_i(t+i-T))\right] \\
&= u_T\left(c_T(t)\right)
\end{aligned}
$$

One clearly sees the effect of a finite life: when an agent ages the number of summands in his utility specification is gradually reduced. Let for instance be $T = 100$ and consider utility of an 99 year old agent in t: this would be given by $U_{T-1}(t)$ and consists of two summands, his current utility at the age of 99 and his discounted expected utility as a 100-year old in the next period. A period later, the agent's decision is based only on the remaining $U_T(t+1) = u_T\left(c_T(t+1)\right)$.
Note that the general specification of (3.4.1) allows for changes in period utility $u_a\left(c_a(t)\right)$ over the life cycle. For instance one could imagine a change in the risk tolerance when an agent matures. Other life-cycle patterns may be introduced by the internal habit formulations like that of Constantinides [1990].

Logarithmic Utility

The general the above model is, the intractable it becomes for a formal analysis. Therefore three simplifications are made.
The first has already been explained in Sect. 3.1 and is the consideration of "only" three cohorts in the overlapping generations model. As birth is at age 0 this translates into $T = 2$ giving each agent a three-period life. With a life expectancy of 90 years as mentioned earlier, each time period covers thus 30

years and the subjective discount factor also corresponds to this period length. The complexity seen in the full display of (3.4.1) makes the reduction to just three generations understandable. Referring to Sect. 3.1.1, it must be emphasized that *three* is the minimum number of cohorts that allows to capture the desired features of uncertainty and labor income participation of the retired. The second simplifying assumption is of more limiting nature. It will be assumed that period utility $u_a(c_a(t))$ is not only the same over the life cycle but also of the logarithmic form:

$$u_a(c_a(t)) = \ln(c_a(t)) \quad \text{for} \quad a = 0, 1, 2. \tag{3.4.2}$$

First, note that (3.4.2) is the limit of the more general power utility form, $u_a(c_a(t)) = (1-\gamma)^{-1} c_a(t)^{1-\gamma}$, for the coefficient of relative risk aversion γ approaching one.[21] Power utility is well established in finance and thus is log utility. From the view of intertemporal consumption-saving models the latter unfortunately gives rise to the feature of myopia. Shown by Merton [1969*b*] and Samuelson [1969], this eliminates all hedging demand with respect to the economy's state variable and makes investors act like in one-period models while the model here is clearly of a long-run perspective.[22] This is a clear drawback of the assumption of (3.4.2). Yet, even without the intertemporal implications the model will still be complex. Furthermore, the general equilibrium implications of the life-cycle aspect, which are not present in the mentioned contributions, eliminates myopic behavior in the sense that other agents' myopic decisions must be addressed when deriving the optimal decision. Finally, even with power utility, agents cannot ignore the fact that they live only three periods. Instead, they adjust their saving and investment decision to their corresponding stage in the life-cycle.

The third simplification follows out of convenience and goes in line with the model's specification so far: instead of using the age dependent indexation resulting from the generality of (3.4.1) the intuitive abbreviations "y", "m" and "o" will be applied.

Using all these assumptions, utility for the young is finally specified as

$$\begin{aligned} U_y(t) &= \ln c_y(t) + \mathbb{E}_t[\beta \ln c_m(t+1) + \beta^2 \ln c_o(t+2)] \\ &= \ln c_y(t) + \mathbb{E}_t[\beta\, U_m(t+1)]\,. \end{aligned} \tag{3.4.3}$$

The middle-aged generation's utility is

$$\begin{aligned} U_m(t) &= \ln c_m(t) + \mathbb{E}_t[\beta \ln c_o(t+1)] \\ &= \ln c_m(t) + \mathbb{E}_t[\beta\, U_o(t+1)], \end{aligned} \tag{3.4.4}$$

and the old cohort's utility is represented by

[21] Hence, log utility satisfies all consistency requirements concerning utility functions as explained for instance in Gollier [2001]. Furthermore, it prevents consumption approaching zero as utility derived from it becomes increasingly negative and is not defined for negative values of consumption.

[22] See also Campbell and Viceira [2001*a*, p. 35] for this interpretation.

$$U_o(t) = \ln c_o(t). \tag{3.4.5}$$

Equations (3.4.3) to (3.4.5) are clearly more tractable than the rolled-out (3.4.1). The different utility rationales for each generation are the basis for the heterogeneity of agents.

3.4.2 Budget Constraints

Having established the objective function for each representative agent, this section will explain the constraints for their optimization.

Timing and Information

Before doing so, note that consumption and investment policies will be such that decisions made in period t can *only* depend on information available in that period. As described in Sect. 3.1.2, the realization of the stochastics in the macroeconomic production function, $\tilde{A}(t)$, $\tilde{\theta}(t)$ and $\tilde{N}_y(t)$, becomes known to all agents at the beginning of each period. To put it formally, there is an increasing sequence of information sets, $\{\mathfrak{I}_t : t = 0, 1, \ldots\}$, underlying the economy and available to all representative agents in period t. $\mathfrak{I}_t$ contains the stochastic total factor productivity and the random production elasticity of real capital in the consumption technology as well as the uncertain size of the young cohort and all their histories up to period t. It also includes all prior investments and savings decisions, real and capital market developments, i.e. up to period $t-1$. Again, it is noteworthy that all agents are perfectly homogeneous with respect to information because they have access to exactly the same information set.[23]

Income Sources

Due to the general equilibrium nature of the model, one would expect labor as well as real capital income to enter the budget constraint. However, the introduction of the human capital security in Sect. 3.3.2 transforms income derived from labor into income from the capital market. It is convenient to assume that members of the young generation are endowed with their gross value of human capital at birth,[24] i.e. with the current market price for the capitalized wage income they earn in their working age. This implies, however, that when middle-aged they do *not* receive wages even though working.

---

[23] An interesting alteration would be the introduction of imperfect information. Such a deviation from perfect knowledge would also allow for learning over the life-cycle. Again, both the economics of information or the addition of behaviouralistic elements are interesting subjects on their own but clearly out of the scope of the model.

[24] The distinction between gross and net value of human capital is only relevant when the labor decision is endogenous as will be seen in Sect. 5.4.2.

Instead, the income to labor is distributed to the agents in form of payoffs on the human capital security. This gives the owners – who will not necessarily be completely identical with the labor suppliers – a return on holding that security.
This finding is key to the model. The assumed tradability of human capital allows to circumvent the non-linearity problems in the budget constraint typically introduced by labor income, which was explained in Sect. 2.2.2. With the human capital security claims on all income sources are traded on the capital market, i.e. represent assets in the classic sense of finance. In contrast to the labor-caused non-linearity problem shown in (2.2.4), with all assets marketable the problem of the budget constraint can be formally expressed in the standard form of an accumulation equation of wealth as used by Merton [1969*b*], Samuelson [1969] or Fama [1970*b*].

Wealth Definition

Having established the timing of events and explained that tradable human capital allows the usage of the standard dynamic accumulation for wealth, this equation can finally be specified. Let $w_a(t)$ be the wealth of an agent aged a, which can be used either for the agent's consumption or his savings, $s_a(t)$:[25]

$$w_a(t) = c_a(t) + s_a(t). \tag{3.4.6}$$

In the portfolio selection problem, the agent must allocate these savings to the different securities of the capital market. As standard in finance, let $\alpha_{K,a}(t)$ be the fraction of savings by an a-aged agent allocated to the physical capital security, $\alpha_{H,a}(t)$ to the human capital security and $1-\alpha_{K,a}(t)-\alpha_{H,a}(t)$ to the riskless security, so that portfolio allocations add up to one. The intertemporal evolvement of the agent's budget constraint is then given by

$$\tilde{w}_{a+1}(t+1) = s_a(t) \times \left\{ \begin{array}{l} \alpha_{K,a}(t) \times \left[\tilde{R}_K(t+1) - R(t)\right] \\ + \alpha_{H,a}(t) \times \left[\tilde{R}_H(t+1) - R(t)\right] \\ + R(t) \end{array} \right\}. \tag{3.4.7}$$

This specification is consistent with the assumptions on timing and information. When an agent decides on his portfolio allocations, the outcome of these investment is unknown as the returns on real and human capital are risky. At the beginning of the next period $\tilde{A}(t+1)$ and $\tilde{\theta}(t+1)$ materialize and – via (3.2.1), (3.2.14) and (3.3.9) – determine the returns and thus the agent's wealth of that period, which is again available for either consumption or saving.

[25] Note that endowing agents with their gross value of human capital at birth and the assumptions of log-utility prevents consumption becoming negative.

Remembering the assumption that the young generation is endowed its gross value of human capital at birth and applying the more appealing y, m, o-notation (3.4.7), reads for the different cohorts as follows:

$$w_y(t) = p_H(t), \tag{3.4.8}$$

$$\tilde{w}_m(t+1) = s_y(t) \times \left\{ \begin{array}{l} \alpha_{K,y}(t) \times \left[\tilde{R}_K(t+1) - R(t)\right] \\ + \alpha_{H,y}(t) \times \left[\tilde{R}_H(t+1) - R(t)\right] \\ + R(t) \end{array} \right\}, \tag{3.4.9}$$

$$\text{and} \quad \tilde{w}_o(t+1) = s_m(t) \times \left\{ \begin{array}{l} \alpha_{K,m}(t) \times \left[\tilde{R}_K(t+1) - R(t)\right] \\ + \alpha_{H,m}(t) \times \left[\tilde{R}_H(t+1) - R(t)\right] \\ + R(t) \end{array} \right\}. \tag{3.4.10}$$

Equations (3.4.8) to (3.4.10) are the intertemporal budget constraints corresponding to the utility maximization problems of (3.4.3) to (3.4.5). Each agent's consumption-savings decision and portfolio selection problem clearly appear in these equations.

3.4.3 Optimal Policies

In order to solve these optimization problems, a simple application of stochastic dynamic programming approach is applied. While this section follows mostly Merton [1983, Appendix], more general explanations on stochastic optimal control can be found in Fleming and Rishel [1975], Malliaris and Brock [1982], Kamien and Schwartz [1991], or Sethi and Thompson [2000].

Definition of Policies

Following straightforwardly from the explanations so far, the consumption and investment policies are such that decisions made in period t depend only on information available in that period. To put it formally, a decision on consumption and portfolio allocations of an agent born in period t is defined as the collection of the $\mathfrak{I}_t$-measurable $(c_y(t), \alpha_{K,y}(t), \alpha_{H,y}(t))$, the $\mathfrak{I}_{t+1}$-measurable $(c_m(t+1), \alpha_{K,m}(t+1), \alpha_{H,m}(t+1))$ and the $\mathfrak{I}_{t+2}$-measurable $c_o(t+2)$. The corresponding savings decisions are always implied by (3.4.6) and the allocation to the riskfree asset is $1 - \alpha_{K,a}(t) - \alpha_{H,a}(t)$ and likewise for $t+1$.

Policy of the Old

Applying the method of stochastic dynamic programming, define indirect utility of an agents as

$$J_a[w_a(t), t] \equiv \max\{U_a(t)\}, \tag{3.4.11}$$

where $w_a(t)$ is the agent's wealth at age a and $U_a(t)$ the lifetime utility of consumption function defined in (3.4.3), (3.4.4) and (3.4.5).
Starting the backward induction of the stochastic dynamic programming approach, it is easily seen from (3.4.5) that the optimal policy of an old agent is given by

$$c_o^*(t) = w_o(t) \tag{3.4.12}$$

where the asterisk denotes optimality. Interpretation of (3.4.12) is straightforward: due to the no-bequest assumption of Sect. 3.4.1, it is optimal for an agent of the old cohort to consume his *entire* wealth as he is in the final period of his life-cycle with certainty. Consequently an old agent does not save, as can be seen from (3.4.6):

$$s_o^*(t) = 0 \tag{3.4.13}$$

Hence, from (3.4.5) and (3.4.11) indirect utility of an old agent is

$$J_o[w_o(t), t] = \max_{c_o(t)}\{\ln c_o(t)\} = \ln c_o^*(t) = \ln w_o(t). \tag{3.4.14}$$

Policy of the Middle

Having solved the optimization problem for the old, backward induction continues with finding the policy of the middle-aged. Indirect utility of such an agent is given through (3.4.4) and (3.4.11) as

$$J_m[w_m(t), t] = \max_{\substack{c_m(t)\\ \alpha_{K,m}(t)\\ \alpha_{H,m}(t)}} \left\{ \ln c_m(t) + \mathbb{E}_t \left[\beta\, U_o(t+1) \right] \right\}. \tag{3.4.15}$$

Applying the result of (3.4.14) by stochastic dynamic programming and using (3.4.10) as well as (3.4.6), this can be rewritten as[26]

$$J_m[w_m(t), t] =$$
$$\max_{\substack{c_m(t)\\ \alpha_{K,m}(t)\\ \alpha_{H,m}(t)}} \left\{ \begin{array}{l} \ln c_m(t) + \\ \beta \mathbb{E}_t \left[\ln \left(\left[\begin{array}{c} w_m(t) \\ - c_m(t) \end{array} \right] \left\{ \begin{array}{c} \alpha_{K,m}(t)[\tilde{R}_K(t+1) - R(t)] \\ + \alpha_{H,m}(t)[\tilde{R}_H(t+1) - R(t)] \\ + R(t) \end{array} \right\} \right) \right] \end{array} \right\}. \tag{3.4.16}$$

In contrast to an old agent, a member of the middle-aged cohort must establish an explicit consumption and an investment policy. Thus he controls $c_m(t)$, $\alpha_{H,m}$ and $\alpha_{K,m}$ in the above equation. The first order conditions of this maximization problem are given by

[26]See App. B.1.1 in case of doubts.

$$\frac{1}{c_m(t)} + \beta\,\mathbb{E}_t\left[\frac{-1}{w_m(t) - c_m(t)}\right] = 0 \tag{3.4.17}$$

and $$\beta\,\mathbb{E}_t\left[\frac{\tilde{R}_i(t+1) - R(t)}{\left\{\begin{array}{l}\alpha_{K,m}(t)[\tilde{R}_K(t+1) - R(t)] + \\ \alpha_{H,m}(t)[\tilde{R}_H(t+1) - R(t)] + R(t)\end{array}\right\}}\right] = 0 \quad \text{for } i = H, K. \tag{3.4.18}$$

As all terms in the expectations of (3.4.17) are known, it immediately gives the optimal consumption policy

$$c_m^*(t) = \frac{1}{\beta + 1} w_m(t)\,. \tag{3.4.19}$$

And by (3.4.6) optimal savings are

$$\begin{aligned} s_m^*(t) &= w_m(t) - c_m^*(t) = w_m(t) - \frac{1}{\beta+1} w_m(t) \\ &= \frac{\beta}{\beta+1} w_m(t)\,. \end{aligned} \tag{3.4.20}$$

Note that second order conditions are satisfied as period utility is strictly concave in consumption. The result is a well-known feature of logarithmic utility: agents consume, respectively save, a deterministic fraction of wealth. Furthermore, the consumption policy is independent from the portfolio policy.[27] Equation set (3.4.18) cannot be further solved for the portfolio fractions without additional informtation. However, once established, the optimal portfolio policy allows a middle-aged agent to transfer his savings from middle- to old-age through an optimized portfolio. Let $\tilde{R}^*_{P,m}(t+1)$ be the return on this portfolio. Obviously it is given by

$$\begin{aligned} \tilde{R}^*_{P,m}(t+1) = \quad & \alpha^*_{K,m}(t)[\tilde{R}_K(t+1) - R(t)] \\ & + \alpha^*_{H,m}(t)[\tilde{R}_H(t+1) - R(t)] + R(t)\,. \end{aligned} \tag{3.4.21}$$

Using this portfolio return the first order conditions of (3.4.18) reads

$$\mathbb{E}_t\left[\frac{\tilde{R}_i(t+1) - R(t)}{\tilde{R}^*_{P,m}(t+1)}\right] = 0 \qquad \text{for } i = H, K. \tag{3.4.22}$$

Furthermore, the portfolio return allows to write the middle's indirect utility as[28]

$$\begin{aligned} J_m[w_m(t), t] &= \beta \ln \beta - (\beta + 1)\ln(\beta + 1) \\ &\quad + (\beta + 1)\ln w_m(t) + \beta\,\mathbb{E}_t\left[\ln \tilde{R}^*_{P,m}(t+1)\right]\,. \end{aligned} \tag{3.4.23}$$

[27] See for instance Ingersoll [1987, pp. 238-240].

[28] The few manipulations are again given in App. B.1.1.

Policy of the Young

The J-function of the middle-aged enables continuation of the backward induction. Indirect utility for a young agent as defined by (3.4.4) and (3.4.11) is

$$J_y[w_y(t), t] = \max_{\substack{c_y(t)\\ \alpha_{K,y}(t)\\ \alpha_{H,y}(t)}} \left\{ \ln c_y(t) + \mathbb{E}_t \left[\beta U_m(t+1) \right] \right\}. \tag{3.4.24}$$

After substituting (3.4.23) and some manipulations, this can be written as[29]

$$J_y[w_y(t), t] = \max_{\substack{c_y(t)\\ \alpha_{K,y}(t)\\ \alpha_{H,y}(t)}} \left\{ \begin{array}{l} \ln c_y(t) - \beta(\beta+1)\ln(\beta+1) + \beta^2 \ln\beta \\ + \beta^2 \mathbb{E}_t \left[\ln \tilde{R}^*_{P,m}(t+2) \right] + \beta(\beta+1) \times \\ \mathbb{E}_t \left[\ln \left(\begin{bmatrix} w_y(t) \\ - c_y(t) \end{bmatrix} \times \left\{ \begin{array}{l} \alpha_{K,y}(t)[\tilde{R}_K(t+1) - R(t)] \\ + \alpha_{H,y}(t)[\tilde{R}_H(t+1) - R(t)] \\ + R(t) \end{array} \right\} \right) \right] \end{array} \right\}. \tag{3.4.25}$$

Like a middle-aged a young agent must establish an actual consumption and investment policy by controlling $c_y(t)$, $\alpha_{H,y}$ and $\alpha_{K,y}$. First order conditions of the optimization are given by

$$\frac{1}{c_y(t)} + \beta(\beta+1)\mathbb{E}_t \left[\frac{-1}{w_y(t) - c_y(t)} \right] = 0 \tag{3.4.26}$$

$$\beta(\beta+1)\mathbb{E}_t \left[\frac{\tilde{R}_i(t+1) - R(t)}{\left\{ \begin{array}{l} \alpha_{K,y}(t)[\tilde{R}_K(t+1) - R(t)] + \\ \alpha_{H,y}(t)[\tilde{R}_H(t+1) - R(t)] + R(t) \end{array} \right\}} \right] = 0 \quad \text{for } i = H, K. \tag{3.4.27}$$

Again, one can easily solve (3.4.26) for the optimal consumption policy. It is obviously given by

$$c^*_y(t) = \frac{1}{\beta^2 + \beta + 1} w_y(t). \tag{3.4.28}$$

Optimal savings are then by application of (3.4.6)

$$\begin{aligned} s^*_y(t) &= w_y(t) - c^*_y(t) = w_y(t) - \frac{1}{\beta^2 + \beta + 1} w_y(t) \\ &= \frac{\beta^2 + \beta}{\beta^2 + \beta + 1} w_y(t). \end{aligned} \tag{3.4.29}$$

[29]See App. B.1.2 for the details.

Like in the case of the middle-aged the equation set (3.4.27) is not further solvable. Still, let it be rewritten by the usage of the return on the optimal portfolio,

$$\begin{aligned}\tilde{R}^*_{P,y}(t+1) = \; & \alpha^*_{K,y}(t)[\tilde{R}_K(t+1) - R(t)] \\ & + \alpha^*_{H,y}(t)[\tilde{R}_H(t+1) - R(t)] + R(t),\end{aligned} \tag{3.4.30}$$

again as

$$\mathbb{E}_t\left[\frac{\tilde{R}_i(t+1) - R(t)}{\tilde{R}^*_{P,y}(t+1)}\right] = 0 \qquad \text{for } i = H, K. \tag{3.4.31}$$

The indirect utility function can now be stated as

$$J_y[w_y(t), t] = \left\{\begin{array}{l}\beta^2 \ln\beta + \beta(\beta+1)(\ln\beta - 1)\ln(\beta+1) \\ -(\beta^2+\beta+1)\ln(\beta^2+\beta+1) + (\beta^2+\beta+1)\ln w_y(t) \\ +\beta(\beta+1)\mathbb{E}_t\left[\ln \tilde{R}^*_{P,y}(t+1)\right] + \beta^2\mathbb{E}_t\left[\ln\tilde{R}^*_{P,m}(t+2)\right]\end{array}\right\} \tag{3.4.32}$$

which is shown in App. B.1.2.

3.4.4 Stochastic Discount Factor

Before proceeding with a summary and the numerical example, this section will give an alternate representation of the agents' problems in form of the stochastic discount factor approach. As explained in Sect. 2.2.3, the SDF allows an elegant and short formalization of an investor's decision rationale.

Derivation

As shown in that section, the stochastic discount factor reflects the agent's marginal rate of substitution with respect to consumption in period $t+1$ for consumption in t under uncertainty. First, the notation of Sect. 2.2.3 must be adjusted to the requirements of the overlapping generations framework. Therefore, let $\tilde{M}_a(t+1)$ be this marginal rate of substitution of an a-aged agent. Applying the derivation with the definitions of lifetime and period utility, (3.4.1) and (3.4.2), the SDF for a young agent can be specified as

$$\begin{aligned}\tilde{M}_y(t+1) &\equiv \frac{\frac{\partial U_y}{\partial \tilde{c}_m(t+1)}}{\frac{\partial U_y}{\partial c_y(t)}} = \beta\frac{\frac{\partial u_m}{\partial \tilde{c}_m(t+1)}}{\frac{\partial u_y}{\partial c_y(t)}} = \beta\frac{\frac{1}{\tilde{c}^*_m(t+1)}}{\frac{1}{c^*_y(t)}} \\ &= \beta\frac{c^*_y(t)}{\tilde{c}^*_m(t+1)}\,.\end{aligned} \tag{3.4.33}$$

Note the usage of optimal consumption policies. This is direct consequence of the optimality assumption at the margin: additional wealth would be consumed otherwise the consumption-saving decisions in each period could not be optimal per se!
Analogously to (3.4.33), the SDF for a middle-aged agent is given by

$$\begin{aligned}\tilde{M}_m(t+1) &\equiv \frac{\frac{\partial U_m}{\partial \tilde{c}_o(t+1)}}{\frac{\partial U_m}{\partial c_m(t)}} = \beta \times \frac{\frac{\partial u_o}{\partial \tilde{c}_o(t+1)}}{\frac{\partial u_m}{\partial c_m(t)}} = \beta \frac{\frac{1}{\tilde{c}_o^*(t+1)}}{\frac{1}{c_m^*(t)}} \\ &= \beta \frac{c_m^*(t)}{\tilde{c}_o^*(t+1)} . \end{aligned} \tag{3.4.34}$$

According to Sect. 3.4.3, the old agent is not faced with an intertemporal decision problem and hence the SDF-approach is not applicable for him.
Drawing on the definitions and results of Sect. 3.4.3, one can further solve $\tilde{M}_y(t+1)$ and $\tilde{M}_m(t+1)$. For the young agent's SDF, start with (3.4.33) and use the result of (3.4.19) to replace $\tilde{c}_m^*(t+1)$:

$$\tilde{M}_y(t+1) = \beta \frac{c_y^*(t)}{\tilde{c}_m^*(t+1)} = \beta \frac{c_y^*(t)}{\frac{1}{\beta+1}\tilde{w}_m(t+1)} .$$

Now plug in (3.4.28) for $c_y^*(t)$ as well as (3.4.9) for $\tilde{w}_m(t+1)$. Since the marginal principle implies optimality, one can use the definition of $\tilde{R}_{P,y}^*(t+1)$ in the latter. Finally, usage of (3.4.29) to replace $s_y^*(t)$ and simplification yields

$$\begin{aligned}\tilde{M}_y(t+1) &= \beta \frac{\frac{1}{\beta^2+\beta+1} w_y(t)}{\frac{1}{\beta+1} s_y^*(t) \tilde{R}_{P,y}^*(t+1)} = \beta \frac{\frac{1}{\beta^2+\beta+1} w_y(t)}{\frac{1}{\beta+1} \frac{\beta^2+\beta}{\beta^2+\beta+1} w_y(t) \tilde{R}_{P,y}^*(t+1)} \\ &= \frac{1}{\tilde{R}_{P,y}^*(t+1)} . \end{aligned} \tag{3.4.35}$$

The discount factor corresponds to the return on the young's optimal portfolio! The simple the result, the unsurprising it is: there are three assets available for postponing consumption. A rational investor takes all of them into consideration, forms a portfolio and uses the resulting discount rate for assessing future consumption options.
This corresponds to the finding of Sect. 2.2.1. In the deterministic setting of the Fisher [1930] model, the gross interest rate is determined by the aggregated optimal intertemporal consumption decisions. This interest rate is simultaneously used as input in the discounting procedure of future endowments. The result of (3.4.35) translates this argument into a setting under uncertainty similar to the illustration in App. A.1. Furthermore, there are three mediums for shifting consumption intertemporal: participation in future profits and wages via the real capital or the human capital security or the riskless asset. At the margin, an agent will optimally combine these possibilities which results in a single portfolio return. This return integrates the consequences of

giving up one consumption unit in the current period and earning the merits of this investment in the subsequent period into a single gross return number. Consequently, this number serves as appropriate discount factor for the next period. Because the returns on human and physical capital are risky, so is the portfolio return and thus the discount factor. In other words it is stochastic. Similar simplifications for the middle-aged agent's SDF lead to[30]

$$\tilde{M}_m(t+1) = \frac{1}{\tilde{R}^*_{P,m}(t+1)}. \tag{3.4.36}$$

The discount factor of a middle-aged corresponds to the return on his optimal portfolio for the same reasoning as for the young agent.
An alternative way to derive the discount factors is based on the envelope condition of the Bellman principle using the indirect utility functions. This is shown in App. B.1.3.

Application

Having established the stochastic discount factors for a young and middle-aged agent, one can apply the above results to restate the decision problems. As the first order condition with respect to consumption of a young, (3.4.26), holds at optimality, it can be written using $\tilde{M}_y(t+1)$:[31]

$$\begin{aligned}
\frac{1}{c_y(t)} &= \beta(\beta+1)\mathbb{E}_t\left[\frac{1}{w_y(t)-c_y(t)}\right] \\
1 &= \mathbb{E}_t\left[\beta\frac{(\beta+1)c^*_y(t)}{w_y(t)-c^*_y(t)}\right] = \mathbb{E}_t\left[\beta\frac{(\beta+1)c^*_y(t)}{s^*_y(t)}\frac{\tilde{R}^*_{P,y}(t+1)}{\tilde{R}^*_{P,y}(t+1)}\right] \\
1 &= \mathbb{E}_t\left[\beta\frac{c^*_y(t)}{\frac{1}{\beta+1}w_m(t+1)}\tilde{R}^*_{P,y}(t+1)\right] = \mathbb{E}_t\left[\beta\frac{c^*_y(t)}{c^*_m(t+1)}\tilde{R}^*_{P,y}(t+1)\right] \\
1 &= \mathbb{E}_t\left[\tilde{M}_y(t+1)\times\tilde{R}^*_{P,y}(t+1)\right].
\end{aligned} \tag{3.4.37}$$

Using (3.4.35) in the first order conditions concerning the portfolio allocations, (3.4.27), reads

$$\mathbb{E}_t\left[\tilde{M}_y(t+1)\left(\tilde{R}_i(t+1)-R(t)\right)\right] = 0 \qquad \text{for } i = H, K. \tag{3.4.38}$$

[30] See App. B.1.3 for details.

[31] Of course, the result follows also trivially from (3.4.35) by

$$1 = \mathbb{E}_t[1] = \mathbb{E}_t\left[\frac{\tilde{R}^*_{P,y}(t+1)}{\tilde{R}^*_{P,y}(t+1)}\right] = \mathbb{E}_t\left[\tilde{M}_y(t+1)\times\tilde{R}^*_{P,y}(t+1)\right],$$

but the derivation from the first order condition is economically more insightful.

Equations (3.4.37) and (3.4.38) are well established results in asset pricing. Extending the explanations of Sect. 2.2.3, they represent *the* central valuations equations. The former relates to the absolute return of an investment – here the agent's portfolio. The latter puts a condition on the excess return of the real capital security, $\tilde{R}_K(t+1) - R(t)$, and of the human capital security, $\tilde{R}_H(t+1) - R(t)$.
Of course analogous valuation equations hold for the middle aged

$$\mathbb{E}_t\left[\tilde{M}_m(t+1) \times \tilde{R}^*_{P,m}(t+1)\right] = 1 \tag{3.4.39}$$

$$\mathbb{E}_t\left[\tilde{M}_m(t+1)\left(\tilde{R}_i(t+1) - R(t)\right)\right] = 0 \qquad \text{for } i = H, K. \tag{3.4.40}$$

3.4.5 Summary of Policies

In this section, the consumption and the portfolio problem of the agents has been presented. While an explicit solution to the decision concerning consumption versus saving has been found for a young, middle-aged and old agent, the corresponding optimal portfolio allocations have not been derived explicitly yet. All the decision rationales have also been presented in the stochastic discount factor approach. Table 3.2 summarizes these findings.
Finally, note that the agents' decisions do not directly depend on ϑ. As shown in Sect. 3.3 this parameter of the capital adjustment technology does influence the individual securities' returns. The agents who are saving and investing in these securities, i.e. the young and the middle-aged, adjust their portfolio allocations based on the decision rationales summarized in Table 3.2. Consequently, the limiting case of fully deprecating real capital must not be investigated separately.

3.4.6 Illustrative Example Economy

With no differences concerning real capital accumulation, one can directly continue the illustration established in Sects. 3.2.6 and 3.3.6.
To extend the numerical example, assume a middle-aged agent possesses wealth of $w_m(0) = \$8.69$ and an old agent $w_o(0) = \$14.25$.[32] According to (3.4.8) the young's wealth is given by the price of the human capital security, thus $w_y(0) = p_H(0) = \$4.06$ according to Sect. 3.3.6. Furthermore let $\beta = 0.6$. As this refers to approximately 30 years this subjective time discount factor corresponds approximately to a standard 0.98 per annum. Based on the results of Sect. 3.4.3, the consumption-saving decisions of the agents can be determined easily:

[32] At first sight, it might seem unrealistic that an old agent's wealth is higher then that of a middle-aged. However, remember that wealth is measured at the beginning of the period; i.e. w_m is before the working period and w_o after. This interpretation makes the assumption more intuitive. Furthermore, bequests which are ignored here, would imply a certain wealth accumulation throughout the life-cycle.

Table 3.2. Summary of agents' decision rationales

		Young Agent	
Wealth	$w_y(t)$	$= p_H(t)$	(3.4.8)
Consumption	$c_y^*(t)$	$= (\beta^2+\beta+1)^{-1}w_y(t)$	(3.4.28)
Saving	$s_y^*(t)$	$= (\beta^2+\beta)(\beta^2+\beta+1)^{-1}w_y(t)$	(3.4.29)
Portfolio	0	$= \mathbb{E}_t\left[\dfrac{\tilde{R}_i(t+1)-R(t)}{\{\alpha_{K,y}^*(t)[\tilde{R}_K(t+1)-R(t)] + \alpha_{H,y}^*(t)[\tilde{R}_H(t+1)-R(t)] + R(t)\}}\right]$	(3.4.27)
SDF	$\tilde{M}_y(t+1)$	$= \tilde{R}_{P,y}^*(t+1)^{-1}$	(3.4.35)
	1	$= \mathbb{E}_t\left[\tilde{M}_y(t+1)\tilde{R}_{P,y}^*(t+1)\right]$	(3.4.37)
	0	$= \mathbb{E}_t\left[\tilde{M}_y(t+1)(\tilde{R}_i(t+1)-R(t))\right]$	(3.4.38)
		Middle-aged Agent	
Wealth	$w_m(t)$	$= s_y^*(t-1)R_{P,y}^*(t)$	(3.4.9)
Consumption	$c_m^*(t)$	$= (\beta+1)^{-1}w_m(t)$	(3.4.19)
Saving	$s_m^*(t)$	$= \beta(\beta+1)^{-1}w_m(t)$	(3.4.20)
Portfolio	0	$= \mathbb{E}_t\left[\dfrac{\tilde{R}_i(t+1)-R(t)}{\{\alpha_{K,m}^*(t)[\tilde{R}_K(t+1)-R(t)] + \alpha_{H,m}^*(t)[\tilde{R}_H(t+1)-R(t)] + R(t)\}}\right]$	(3.4.18)
SDF	$\tilde{M}_m(t+1)$	$= \tilde{R}_{P,m}^*(t+1)^{-1}$	(3.4.36)
	1	$= \mathbb{E}_t\left[\tilde{M}_m(t+1)\tilde{R}_{P,m}^*(t+1)\right]$	(3.4.40)
	0	$= \mathbb{E}_t\left[\tilde{M}_m(t+1)(\tilde{R}_i(t+1)-R(t))\right]$	(3.4.40)
		Old Agent	
Wealth	$w_o(t)$	$= s_m^*(t-1)R_{P,m}^*(t)$	(3.4.10)
Consumption	$c_o(t)^*$	$= w_o(t)$	(3.4.12)
Saving	$s_o(t)^*$	$= 0$	(3.4.13)
Portfolio and SDF		Not applicable	

Remarks: The table summarizes the results on the consumption and on the portfolio problem, which has not yet been explicitly solved. For each agent $i = H, K$.

$$\begin{aligned} c_y^*(0) &= \frac{1}{0.6^2+0.6+1}\$4.06 \approx \$2.07 & s_y^*(0) &= \frac{0.6^2+0.6}{0.6^2+0.6+1}\$4.06 \approx \$1.99 \\ c_m^*(0) &= \frac{1}{0.6+1}\$8.69 \approx \$5.43 & s_m^*(0) &= \frac{0.6}{0.6+1}\$8.69 \approx \$3.26 \\ c_o^*(0) &= \$14.25 & s_o^*(0) &= \$0.00. \end{aligned}$$

The actual portfolio rules have not been found yet. Still, the allocation rule implies that the young agent chooses portfolio weights, $\alpha_{h,y}^*(0)$ and $\alpha_{K,y}^*(0)$, so that it holds for $i = H, K$.

$$\mathrm{E}_0 \left[\frac{\tilde{R}_i(1) - R(0)}{\alpha_{K,y}^*(0)[\tilde{R}_K(1) - R(0)] + \alpha_{H,y}^*(0)[\tilde{R}_H(1) - R(0)] + R(0)} \right] = 0.$$

Let the riskfree rate be 259%. Using the results of Sect. 3.3.6 this means that for the excess return on the human capital security the following must hold

$$\left[\begin{array}{l} \dfrac{226\% - 259\%}{\alpha_{K,y}^*(0)[115\% - 259\%] + \alpha_{H,y}^*(0)[226\% - 259\%] + 259\%} \times P_1 \\ + \dfrac{453\% - 259\%}{\alpha_{K,y}^*(0)[230\% - 259\%] + \alpha_{H,y}^*(0)[453\% - 259\%] + 259\%} \times P_2 \\ + \dfrac{195\% - 259\%}{\alpha_{K,y}^*(0)[264\% - 259\%] + \alpha_{H,y}^*(0)[195\% - 259\%] + 259\%} \times P_3 \\ + \dfrac{390\% - 259\%}{\alpha_{K,y}^*(0)[528\% - 259\%] + \alpha_{H,y}^*(0)[390\% - 259\%] + 259\%} \times P_4 \end{array} \right] = 0.$$

Together with an analogous condition for the real capital security, this gives the optimal portfolio weights. Furthermore, these conditions will also hold for the middle-aged agent. Section 3.5 draws heavily on this last observation.

3.5 Capital Market Equilibrium

Having established the decision rationales for agents of all cohorts, equilibrium in the capital market can be established. Based on the concept of representative agents, the portfolio rules are aggregated across generations to derive equilibrium conditions for the capital market which correspond to well known concepts in financial theory. These conditions allow to solve for the equilibrium portfolio weights and returns. Again the stochastic discount factor approach supplements the explanations.

3.5.1 Aggregating Assumptions

When going from individual agents' optimal behavior to an equilibrium perspective, the concept of representative agents plays a key role.

Representative Agent

As mentioned in Sect. 3.1.1, each cohort can be identified with a representative agent, as all members of a cohort are identical. Therefore, the individual agents modeled so far present not only a particular member of that cohort but serve as a representative of *each* agent in that cohort. This homogeneity assumption allows to summarize all agents of one generation by a single representative agent for that cohort. Note that the assumptions on the labor market imposed in Sect. 3.3.2 exclude heterogeneity among a cohort's agents with respect to labor income. As all agents of a particular generation provide the same amount of labor service, their wages are perfectly correlated. Hence, there is no need to require a mechanism to pool idiosyncratic income shocks *within* a generation since such a idiosyncratic variation does not exist. Instead, the simplicity of the labor market automatically implies the aggregation into a cohort's representative.[33]
However, in contrast to the model of Constantinides et al. [2002] it would be shortsighted not to address the cohort sizes. As their model is based on an exchange economy in spirit of Lucas [1978], neither production nor general equilibrium are addressed. Hence there is no need for demographic variables apart of the distinction of investors with respect to the life-cycle. In the analysis here, however, these aspects must be included as explained in Chap. 2. Therefore, another actual aggregation is applied summarizing the representative agent of the young, the middle-aged and the old generation into a single population representative. In this step, it is natural to scale the generation representative agents with the cohort sizes, addressing their relative importance for the economy, and then summarize the three scaled agents to the population representative. As explained in Constantinides [1989], this aggregation requires completeness of the capital market in order not to affect equilibrium prices or consumption levels. Here, this completeness is implied by the setup of the model: because it is assumed that future labor income is tradable via the human capital security, all cohort-specific shocks are insurable on the capital market. This insurance element is the core feature for replicating completeness with the PAYGO-system.[34] Hence, all generations' endowment and decision variables are integrated into the economy's representative agent, whose composition depends on the demographic structure.[35]

[33]Naturally, it would present an interesting extension to allow for idiosyncratic and uninsurable income shocks. This would then extend the analysis of functional income distribution to include aspects of the personal income dispersion.

[34]Note, however, that completeness does not include the economy-wide shocks $\tilde{A}(t)$. These shocks represent uninsurable "endowment" risk and hence completeness with respect to it is not required for the aggregation. To see this, imagine an Arrow-Debreu economy with a complete market. Scaling all payoffs with the same factor has no effect on completeness.

[35]In essence the result achieved by above two-step procedure could also be achieved by directly integrating across all agents of all generations and giving agents

Aggregating Definitions

Based on the above explanations, let the aggregate variables of a cohort be

$$\begin{aligned} W_y(t) &\equiv N_y(t)w_y(t), & C_y(t) &\equiv N_y(t)c^*_y(t), & S_y(t) &\equiv N_y(t)s^*_y(t), \\ W_m(t) &\equiv N_m(t)w_m(t), & C_m(t) &\equiv N_m(t)c^*_m(t), & S_m(t) &\equiv N_m(t)s^*_m(t), \\ W_o(t) &\equiv N_o(t)w_o(t), & C_o(t) &\equiv N_o(t)c^*_o(t), \end{aligned} \tag{3.5.1}$$

where the the cohort sizes, $N_y(t)$, $N_m(t)$ and $N_o(t)$ have been established in Sect. 3.1.1. Using these definitions, the economy's aggregate variables are then given by

$$\begin{aligned} W(t) &\equiv W_y(t) + W_m(t) + W_o(t), \\ C(t) &\equiv C_y(t) + C_m(t) + C_o(t), \\ \text{and} \quad S(t) &\equiv S_y(t) + S_m(t). \end{aligned} \tag{3.5.2}$$

As the old do not save, their aggregate savings variable can be neglected.

Aggregated Policies

With the definitions of (3.5.1), one can immediately solve for the cohort-wide variables. Based on the results of (3.4.12), (3.4.19) and (3.4.28) as well as (3.4.20) and (3.4.29), it follows that

$$C_y(t) = \frac{1}{\beta^2+\beta+1} W_y(t), \quad C_m(t) = \frac{1}{\beta+1} W_m(t), \quad C_o(t) = W_o(t), \tag{3.5.3}$$

$$S_y(t) = \frac{\beta^2+\beta}{\beta^2+\beta+1} W_y(t), \quad S_m(t) = \frac{\beta}{\beta+1} W_m(t). \tag{3.5.4}$$

Since the above equations mimic perfectly an individual agent's decision rationales, these results reemphasize the representativeness of the cohort-aggregating agent. This can further be seen from combining (3.5.3) and (3.5.4) to establish budget constraints for the cohorts,

$$W_y(t) = C_y(t) + S_y(t) \quad \text{and} \quad W_m(t) = C_m(t) + S_m(t), \tag{3.5.5}$$

which is consistent with taking the budget statement (3.4.6) to the cohort-level via (3.5.1). With the aggregation procedure in place, one can continue to establish equilibrium in the capital market.

3.5.2 Deriving Equilibrium

To achieve equilibrium one first identifies the individual securities' weights in the market portfolio and then derives conditions guaranteeing equilibrium.

---

within a generation the same weights. Above remarks on the completeness would allow to follow Constantinides [1982, Lemma 1]. However, the stepwise aggregation is more intuitive.

Portfolio Rules

As briefly indicated in Sect. 3.4.6, the key insight is the observation that the portfolio rules for the young and the middle-aged are the same.[36] Compare the conditions on which these generations base their decisions. From (3.4.27), these are for the young

$$0 = \mathbb{E}_t \left[\frac{\tilde{R}_H(t+1) - R(t)}{\alpha^*_{K,y}(t)[\tilde{R}_K(t+1) - R(t)] + \alpha^*_{H,y}(t)[\tilde{R}_H(t+1) - R(t)] + R(t)} \right]$$

$$0 = \mathbb{E}_t \left[\frac{\tilde{R}_K(t+1) - R(t)}{\alpha^*_{K,y}(t)[\tilde{R}_K(t+1) - R(t)] + \alpha^*_{H,y}(t)[\tilde{R}_H(t+1) - R(t)] + R(t)} \right]$$

and from (3.4.18) for the middle-aged

$$0 = \mathbb{E}_t \left[\frac{\tilde{R}_H(t+1) - R(t)}{\alpha^*_{K,m}(t)[\tilde{R}_K(t+1) - R(t)] + \alpha^*_{H,m}(t)[\tilde{R}_H(t+1) - R(t)] + R(t)} \right]$$

$$0 = \mathbb{E}_t \left[\frac{\tilde{R}_K(t+1) - R(t)}{\alpha^*_{K,m}(t)[\tilde{R}_K(t+1) - R(t)] + \alpha^*_{H,m}(t)[\tilde{R}_H(t+1) - R(t)] + R(t)} \right] .$$

By inspection of these rules, the optimal portfolios' fractional allocations to the securities, $\alpha_{H,i}$ and $\alpha_{K,i}$, must be identical for the young and the middle-aged!

Nota bene: this identity means that young and middle-aged allocate savings in exactly the *same proportions* across the securities. This does not imply that absolute investments in human and real capital are equal. The overall level invested in each security depends not only on the fractional allocations but also on the absolute savings of each cohort. This difference is also reflected in the intertemporal budget constraints, (3.4.9) and (3.4.10): on the one hand, the consumption versus savings problem – i.e. the decision on the absolute amount of investments in the financial markets – enters via the savings. On the other hand, the allocation decision – i.e. the fractional distribution of the savings across the securities – is represented by the portfolio shares $\alpha_{H,a}(t)$ and $\alpha_{K,a}(t)$.

Market Portfolio

If the optimal values of these allocations are the same for both generations engaged in saving, then this *joint* optimal portfolio is the *market portfolio.* This implies

$$\alpha^M_K(t) = \alpha^*_{K,y}(t) = \alpha^*_{K,m}(t) \quad \text{and} \quad \alpha^M_H(t) = \alpha^*_{H,y}(t) = \alpha^*_{H,m}(t), \quad (3.5.6)$$

[36] Having established the representative agent concept, the terms "the young" and "the middle-aged" obviously refer to the entire cohort, comprising all the individual cohort members.

where $\alpha_H^M(t)$ and $\alpha_K^M(t)$ denote the proportions of the human and physical capital security in the market portfolio. Note how this observation interacts with the aggregating assumption of Sect. 3.5.1: as the decision rationales of both saving-engaged cohort representative agents are identical, it does not matter whether the marginal investor is a young or a middle-aged agent. Therefore, the aggregation into the economy wide agent is easily achieved. One could think of it as identical investors allocating a standardized unit in the capital market. Thereby the number of investors depends on the number of young and middle-aged agents and their respective absolute amount of wealth dedicated to savings.
Using the information on the supply of the securities as given in Sect. 3.3.4 and summarized in Table 3.1 allows to further identify the market portfolio. Since the riskless security is in zero net supply, it holds that

$$\alpha_H^M(t) = 1 - \alpha_K^M(t). \tag{3.5.7}$$

As both prices and number of securities outstanding have been established for human and physical capital, one can calculate the total value of the market and infer the value weighted fractions for each security

$$\alpha_K^M(t) = \frac{K(t+1)p_K(t)}{K(t+1)p_K(t) + N_m(t+1)p_H(t)} \tag{3.5.8}$$

$$\text{and} \quad \alpha_H^M(t) = \frac{N_m(t+1)p_H(t)}{K(t+1)p_K(t) + N_m(t+1)p_H(t)}. \tag{3.5.9}$$

By construction, this means that the market portfolio comprises all securities outstanding in proportion to their market values. Drawing on the results of Sect. 3.2, these weights can further be identified. Using (3.2.12) for the value of the physical capital securities, $K(t+1)p_K(t)$, and remembering that the population assumptions imply $N_m(t+1) = N_y(t)$ allows to restate the weights in the market portfolio as

$$\alpha_K^M(t) = \frac{\frac{1}{\vartheta}I(t)}{\frac{1}{\vartheta}I(t) + N_y(t)p_H(t)} = \frac{I(t)}{I(t) + \vartheta N_y(t)p_H(t)} \tag{3.5.10}$$

$$\text{and} \quad \alpha_H^M(t) = \frac{N_y(t)p_H(t)}{\frac{1}{\vartheta}I(t) + N_y(t)p_H(t)} = \frac{\vartheta N_y(t)p_H(t))}{I(t) + \vartheta N_y(t)p_H(t)}. \tag{3.5.11}$$

Equilibrium Conditions

Inserting these last expressions into the portfolio rules (3.4.18) or (3.4.27) yields well known conditions for the equilibrium. As shown in App. B.2.1 it must hold that

$$\mathbb{E}_t\left[\frac{\tilde{R}_K(t+1) - R(t)}{\tilde{R}_M(t+1)}\right] = 0 \quad \text{and} \quad \mathbb{E}_t\left[\frac{\tilde{R}_H(t+1) - R(t)}{\tilde{R}_M(t+1)}\right] = 0. \tag{3.5.12}$$

The expected excess return on each security, $\tilde{R}_i(t+1) - R(t)$, relative to the return on the capital market is zero. Furthermore, comparing above conditions with (3.4.21) and (3.4.30) one sees that

$$\tilde{R}_M(t+1) = \tilde{R}^*_{P,y}(t+1) = \tilde{R}^*_{P,m}(t+1), \tag{3.5.13}$$

which could also be derived from the findings on the market portfolio's weights. Since all agents take all securities optimally into consideration for their portfolio formations and since in equilibrium the entire market is held, each agent ends up with a portfolio reflecting the market's weights. Rearranging and manipulating (3.5.12) as shown in App. B.2.1 leads to an equally important and equally known result:

$$1 = \mathbb{E}_t\left[\frac{R(t)}{\tilde{R}_M(t+1)}\right] = \mathbb{E}_t\left[\frac{\tilde{R}_K(t+1)}{\tilde{R}_M(t+1)}\right] = \mathbb{E}_t\left[\frac{\tilde{R}_H(t+1)}{\tilde{R}_M(t+1)}\right] \tag{3.5.14}$$

The marginal relative contribution of a security's return to the market portfolio return is one for all securities. This implies that agents split their investments in the different securities in such a way that expected marginal returns relative to the – individual and thus market – portfolio return are equal across *all* securities. Furthermore, at the optimum, shifting a marginal unit of savings from one security into another does not yield any extra return, because the benefit of doing so equals the opportunity costs. In other words, each security is expected to contribute the same risk-adjusted return to the individual portfolios and hence to the market return.

Consequently, agents are only willing to hold a security in equilibrium if its expected return satisfies above equation. This condition is achieved by the usual price mechanisms on the financial markets. For instance, assume the human capital security is currently trading at a too high a price to satisfy (3.5.14), i.e. given its expected payoff the return on the human capital security is too low. Then agents will drive down this security's price by selling it, thereby increasing the expected return until the condition is met. At the same time, however, the expected return on the real capital security will be too high, i.e. its price too low to satisfy (3.5.14). The reason for this is that with only two kinds of securities the return of one must be above the market's return when the other's return is below.[37] Then for the real capital security the opposite will obviously happen: purchasing orders from the agents leads to an appreciation of its price and hence a decline of its expected return. Both adjustments will also influence the aggregated portfolio return,[38] but ultimately (3.5.14) will hold again. Observe that investing in one security implies divesting the other, so that no wealth is withdrawn from the financial markets. It is only a rebalancing procedure until the allocations establish an equilibrium.

[37] This is a consequence of the fact that the market is a weighted average of the individual returns and the weights cannot be negative in an equilibrium.

[38] As such a situation does actually not represent an equilibrium, it is not yet the "market" portfolio per se.

3.5.3 Market Portfolio Weights

From above equilibrium condition the weights of the market portfolio can now be derived involving various steps of simplifications.

Investment-Output Ratio

Take the return on the capital market from (3.3.10) as well as the return on a security, for instance on the human capital security and hence (3.3.9). Plugging them into (3.5.14) gives after some simplifications[39]

$$1 = \mathbb{E}_t\left[\frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{\tilde{Y}(t+1)+\frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}\right]\left(1+\frac{I(t)}{\vartheta N_y(t)p_H(t)}\right). \quad (3.5.15)$$

Now define the ratio of investment to output, $\psi(t) \boxminus I(t)/Y(t)$, and assume that all agents belief that

$$\mathbb{E}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}}\right] = \frac{1}{1+\frac{1-\vartheta}{\vartheta}\bar{\psi}} \quad \text{and} \quad \mathrm{Cov}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}},\tilde{\theta}\right] = 0, \quad (3.5.16)$$

where Cov $[\circ]$ is the covariance.
The assumptions of (3.5.16) reflect and extend the findings of Abel [2003, Proposition 1]: It exists a long run constant investment-output ratio that remains constant even though random shock generate stochastic movements in the ratio of physical capital per worker.[40] As agents have also a constant expectation concerning the random elasticity parameter, i.e. $\mathbb{E}_t[\tilde{\theta}] = \bar{\theta}$ as given in Sect. 3.2.1, Abel's [2003] results are not affected by this additional uncertainty. Of course the investment formation process is more complex, but the saving rules of all generations reflect the myopic property of the utility assumption and do hence *not* depend on any state variables. In other words and inferrable from (3.5.4), saving is linear in the wealth distribution regardless of future return outlook.[41]
As in Abel [2003] even with a constant investment-output ratio the price of physical capital, i.e. the q-ratio, can still vary. Each period, shocks in terms of uncertain total factor productivity and deviations from the expected elasticity parameter affect the economic system and establish a new price of physical capital. Consequently, the commonly made observation that with a perfect capital market the q-ratio can only deviate from unity in the short run is

[39] See App. B.2.2.

[40] Nota bene that $\bar{\psi}$ is *not* this long run ratio, but defined more conveniently for the following derivations. Appendix B.2.2 gives a relation between the actual mean and $\bar{\psi}$.

[41] This is clearly one of the model's major simplifications and does hardly make the savings decision any more endogenous than in other models. Still, considering the time horizon the approximations should be not too unrealistic.

not applicable here. Before the past deviations can be eliminated new shocks have already materialized. Furthermore, Sect. 3.6 will verify that this model's investment to output ratio is indeed independent of time.
Using the assumptions and the investment-output ratio allows to calculate the market capitalization of the human capital in equilibrium as shown in App. B.2.2:

$$N_y(t)p_H(t) = \frac{1-\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} \times I(t). \qquad (3.5.17)$$

Since the population structure implies $N_y(t) = N_m(t+1)$, one can further manipulate above equation. By undoing the usage of (3.2.12), the total market capitalizations of both securities can be related to each other:

$$\begin{aligned} N_y(t)p_H(t) = N_m(t+1)p_H(t) &= \frac{1-\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} \times I(t) \\ &= \frac{(1-\bar{\theta})\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} \times K(t+1)p_K(t). \qquad (3.5.18) \end{aligned}$$

Finally solving this for the total market value yields according to App. B.2.2

$$K(t+1)p_K(t) + N_m(t+1)p_H(t) = \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}\frac{1}{\vartheta}I(t). \qquad (3.5.19)$$

Equilibrium Weights

Using the above results in (3.5.8) and (3.5.9) gives the weights of the market portfolio as

$$\alpha_K^M(t) = \frac{K(t+1)p_K(t)}{\frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}K(t+1)p_K(t)} = \frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta} \qquad (3.5.20)$$

$$\text{and} \quad \alpha_H^M(t) = \frac{\frac{(1-\bar{\theta})\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}K(t+1)p_K(t)}{\frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}K(t+1)p_K(t)} = \frac{(1-\bar{\theta})\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta}. \qquad (3.5.21)$$

Note that the weight of the human capital security in the market portfolio can also be derived from (3.5.7). As $\alpha_H^M(t) = 1 - \alpha_K^M(t)$ it follows

$$\begin{aligned} \alpha_H^M(t) = 1 - \frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta} &= \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta} - \frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta} \\ = \frac{(1-\vartheta)\bar{\psi}+\vartheta-(1-\vartheta)\bar{\psi}-\vartheta\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta} &= \frac{(1-\bar{\theta})\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta}, \end{aligned}$$

which confirms the result of (3.5.21).
The market portfolio weights of the two security classes are constant over time in this setting. This implies that the relative importance of human capital and

physical capital with respect to each other does *not* change. While this result is formally a consequence of the assumptions on the elasticity parameter of the consumption technology, $\bar{\theta}$, and on the ratio of investment to output, $\bar{\psi}$, it reflects the growth-theoretically driven view of balancing importance of labor and real capital as macroeconomic input factors. Because the underlying production technologies exhibit the neoclassical properties of positive but decreasing marginal productiveness, valuation of the economy's input factors is *relative* to each other and driven by the elasticity parameters. Equation (3.5.18) reflects that, if one of the the input factors is zero than the other is worthless as well! As the production technology of (3.2.1) does not allow to completely substitute one input factor by the other, even a very high physical capital stock is worthless if there is no labor force to work on it. Vice versa, a large number of workers cannot produce output without machinery. But note that market weights of human and real capital are *not identical*, as (3.5.20) and (3.5.21) are different in general. They are rather constantly connected via the substitutional possibilities in the economy real side.

3.5.4 Returns and Interest Rate

Having connected the equilibrium market values one can proceed to solve for the individual securities' returns that form the capital market.

Return on Capital Market

Starting with (3.3.10) for the capital market's return and using the investment-output ratio and (3.5.19) from the last section allows to establish the relation for the equilibrium return on the market portfolio:

$$\begin{aligned}\tilde{R}_M(t+1) &= \frac{\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}{K(t+1)p_K(t) + N_m(t+1)p_H(t)} \\ &= \left[(1-\vartheta)\tilde{\psi}(t+1) + \vartheta\right] \frac{(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}}{(1-\vartheta)\bar{\psi} + \vartheta} \frac{\tilde{Y}(t+1)}{I(t)}. \quad (3.5.22)\end{aligned}$$

This equation connects the real and the financial perspectives of the economy. As intuition would imply, the return on the market depends positively on the subsequent's period output: there is more to be distributed when output is higher and the complete capital market captures all rental-earning productive factors. Equally intuitive would be the interpretation that the market's return is negatively related to current investments, as relative abundance of physical capital implies a lower rental. But (3.5.22) shows that $\tilde{Y}(t+1)$ also depends on $I(t)$ according to (3.2.1) and (3.2.7). Hence such a conclusion would be premature. Still, one can claim that a higher subsequent investment-output ratio $\tilde{\psi}(t+1)$ implies a higher return on the capital market ceteris paribus. The reason for this is that the investment-output ratio reflects the valuation of real capital in terms of consumption units. As in models based on the q-Theory of

Tobin [1969], which have been introduced in Sect. 2.2.4, investments will be high when real capital is highly valued. But this implies that already installed physical capital will earn a high rental in the capital adjustment technology as well, which results in a higher return of the market.[42]

Return on Real and Human Capital

In analogy to the above, equilibrium returns on the real and human capital securities can be calculated. Taking the returns as summarized in Table 3.1 and using the investment-output ratio as well as (3.5.19) yields

$$\begin{aligned}\tilde{R}_K(t+1) &= \frac{\tilde{\theta}(t+1)\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}{K(t+1)p_K(t)} \\ &= \frac{\tilde{\theta}(t+1)\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta}\tilde{\psi}(t+1)\tilde{Y}(t+1)}{\frac{1}{\vartheta}I(t)} \\ &= \left[(1-\vartheta)\tilde{\psi}(t+1) + \vartheta\tilde{\theta}(t+1)\right]\frac{\tilde{Y}(t+1)}{I(t)} \end{aligned} \tag{3.5.23}$$

for physical capital and, for human capital,

$$\begin{aligned}\tilde{R}_H(t+1) &= \frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{N_m(t+1)p_H(t)} \\ &= \frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{\frac{1-\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}I(t)} \\ &= \left(1-\tilde{\theta}(t+1)\right)\frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{1-\bar{\theta}}\frac{\tilde{Y}(t+1)}{I(t)}. \end{aligned} \tag{3.5.24}$$

A comparison of (3.5.23) and (3.5.24) results in the insight, that only $\tilde{R}_K(t+1)$ depends on the subsequent investment-output ratio $\tilde{\psi}(t+1)$. This is in line with the q-Theory explanation given for the capital market return. Furthermore, obverse that $\tilde{R}_K(t+1) = \tilde{R}_H(t+1)$ when neither the realizations of the elasticity parameter nor that of the investment-output ratio deviate from their expected values.[43]

This is an interesting and important result: returns on real and human capital are exactly the same if nothing unexpected happens. As (3.5.23) and (3.5.24) refer to equilibrium this also includes any risk premium considerations! When there are no surprises the return on both kinds of capital will be the same. Given the neoclassical setting of the model, this is not surprising. As returns

[42] To show these relations formally, one actually needs to take the respective derivatives and investigate their signs. This exercise is trivial as by assumption $\vartheta \geq 0$ and $1-\vartheta \geq 0$.

[43] To be more accurate it should read, "when the investment-output ratio equals the conveniently defined $\bar{\psi}$," which is related to the expected investment-output ratio as given in Appendix B.2.2.

are related to the marginal productivity of each factor and as neoclassical production implies balancing of them, they must be the same if there is perfect foresight. Any differences in the equilibrium returns will be induced by deviations in the elasticity parameter and the investment-output ratio. The statistical distributions if those, which have not been specified here, determine the risk premia and lead to different returns for real and human capital.

Riskfree Return

Finally, the shadow riskfree rate $R(t)$ can easily be derived. Take the first relation in (3.5.14), retrieve the certain $R(t)$ out of the expectations and reformulate it as

$$R(t) = \frac{1}{\mathbb{E}_t\left[\left(\tilde{R}_M(t+1)\right)^{-1}\right]}. \tag{3.5.25}$$

Using the result for the return on the capital market, (3.5.22), in (3.5.25) leads to

$$\begin{aligned} R(t) &= \frac{1}{\mathbb{E}_t\left[\left(\tilde{R}_M(t+1)\right)^{-1}\right]} \\ &= \frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta}\frac{1}{I(t)}\left(\mathbb{E}_t\left[\frac{1}{\left[(1-\vartheta)\tilde{\psi}(t+1)+\vartheta\right]\tilde{Y}(t+1)}\right]\right)^{-1} \end{aligned} \tag{3.5.26}$$

Since $0 \leq \vartheta \leq 1$ and investments, output and the investment-output ratio are non-negative the riskfree return will be non negative. Shifting (3.5.26) by one period and taking period t's expectations gives

$$\mathbb{E}_t\left[\tilde{R}(t+1)\right] = \frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta}\left(\mathbb{E}_t\left[\frac{\tilde{I}(t+1)}{\left[(1-\vartheta)\tilde{\psi}(t+2)+\vartheta\right]\tilde{Y}(t+2)}\right]\right)^{-1}.$$

That means the specification for the riskless rate could be extended to a simple model for the term structure of real interest rates. Because the periods are of a rather long term, this term structure would apply to long and very long maturities.

3.5.5 SDF Approach

The stochastic discount factor approach introduced in Sect. 3.4.4 gave supplementing insights into the agents' decision problems. This section continues this explanation and shows how above results can be obtained applying the SDF framework.

Equilibrium Conditions

Using for instance the SDF formulation for the young, one can establish the equivalent to (3.5.14). Starting with (3.4.38), extending the product and splitting the expectations one arrives at

$$\mathbb{E}_t\left[\tilde{M}_y(t+1)R(t)\right] = \mathbb{E}_t\left[\tilde{M}_y(t+1)\tilde{R}_H(t+1)\right] = \mathbb{E}_t\left[\tilde{M}_y(t+1)\tilde{R}_K(t+1)\right]. \tag{3.5.27}$$

Then the same procedure as in Appendix B.2.1, i.e. multiplying with the respective market weights, adding up and noting that the market weights sum to one, allows to extend the above relations to the market's return. Finally, combing (3.4.35) and (3.5.13) with this gives the well known

$$\begin{aligned} 1 &= \mathbb{E}_t\left[\tilde{M}_y(t+1)R(t)\right] = \mathbb{E}_t\left[\tilde{M}_y(t+1)\tilde{R}_M(t+1)\right] \\ &= \mathbb{E}_t\left[\tilde{M}_y(t+1)\tilde{R}_K(t+1)\right] = \mathbb{E}_t\left[\tilde{M}_y(t+1)\tilde{R}_H(t+1)\right]. \end{aligned} \tag{3.5.28}$$

This is the Euler equation in the standard SDF-approach. The same relationship can analogously be established with the stochastic discount factor for the middle-aged, $\tilde{M}_m(t+1)$.

Explicit Solution for SDF

Using the results of Sect. 3.5.4, the explicit solutions for the stochastic discount factor can be stated. Plugging (3.5.22) into the (3.4.35) and (3.4.36) reads:

$$\begin{aligned} \tilde{M}_y(t+1) &= \frac{1}{\tilde{R}^*_{P,y}(t+1)} = \frac{1}{\tilde{R}_M(t+1)} \\ &= \frac{1}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\tilde{\psi}(t+1)+\vartheta} \frac{I(t)}{\tilde{Y}(t+1)}, \end{aligned} \tag{3.5.29}$$

$$\begin{aligned} \tilde{M}_m(t+1) &= \frac{1}{\tilde{R}^*_{P,m}(t+1)} = \frac{1}{\tilde{R}_M(t+1)} \\ &= \frac{1}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\tilde{\psi}(t+1)+\vartheta} \frac{I(t)}{\tilde{Y}(t+1)}. \end{aligned} \tag{3.5.30}$$

The connection of the SDFs to the underlying real sector is evident: the higher future output and the lower current investments, the lower is the stochastic discount factor. This is intuitive, since a low discount factor corresponds to a high discount rate, and a high discount rate reflects high returns on the capital market. Extending the interpretation of Fig. 2.2 confirms this: the slope of the line representing the capital market depends positively on the future output and negatively on the forgone consumption – i.e. investments. Equations (3.5.29) and (3.5.30) reflect the same argument, yet not for the discount rate but for the reciprocal discount factor.

While restating all the findings with the SDF-approach might seem unnecessary at first glance, there are three items achieved by doing so. Firstly, using this standard approach, which has become very fashionable with financial economists, consistency of this framework with standard theory of finance is conveniently shown. Secondly, the identity of the stochastic discount factor for the young and the middle-aged, as e.g. seen in (3.5.29) and (3.5.30), formally proves that markets are complete in this setting. As shown in Cochrane [2001, Chap. 4], for instance, only in complete markets a unique stochastic discount factor prices all the assets. In the setting here the possibility to trade human capital allows the working generation to eliminate, respectively share, its risk concerning the labor income so that no sources of idiosyncratic uncertainty remain. This results in identical SDFs for the young and the middle-aged. And thirdly, the model extends the SDF-approach. Usually it is heavily used in consumption-based models. Here, it is shown how the stochastic discount factor and the marginal rate of transformation behave consistently. This reflects the necessity to extend the standard capital market theory for the long-run perspective on the pension problem as discussed in Chap. 2.[44]

3.5.6 Full Depreciation

Continuing Sect. 3.3.5's examination of the degenerate case, this section derives the capital market equilibrium when real capital fully depreciates. While aggregation and the steps towards equilibrium are unaffected by the relevant parameter ϑ, substantial simplifications can be achieved for the market portfolio weights. Taking the limit of the market capitalization of human capital, (3.5.17), for $\vartheta = 1$ yields

$$N_m(t+1)p_H(t)|_{\vartheta=1} = \frac{1-\bar{\theta}}{(1-1)\bar{\psi} + 1 \times \bar{\theta}} \times I(t) = \frac{1-\bar{\theta}}{\bar{\theta}} I(t). \quad (3.5.31)$$

As shown in Sect. 3.2.5 the per-unit price of real capital in absence of adjustment costs, $p_K(t)|_{\vartheta=1}$, is unity. Using this and the degenerate adjustment technology as given in (3.2.8), the market value of the real capital security is given by

$$K(t+1)|_{\vartheta=1}\, p_K(t)|_{\vartheta=1} = I(t)\,. \quad (3.5.32)$$

The same result is obtained by plugging $\vartheta = 1$ into (3.2.12). According to (3.5.20) and (3.5.21), the capital market weights reduce conveniently to

$$\alpha_K^M(t)|_{\vartheta=1} = \frac{(1-1)\bar{\psi} + 1 \times \bar{\theta}}{(1-1)\bar{\psi} + 1} = \bar{\theta}$$

$$\text{and} \quad \alpha_H^M(t)|_{\vartheta=1} = \frac{(1-\bar{\theta}) \times 1}{(1-1)\bar{\psi} + 1} = 1 - \bar{\theta}, \quad (3.5.33)$$

[44]For convenience, note that all equilibrium findings will be summarized in Sect. 3.7.

which obviously sum to one. Naturally, this could have also been obtained easily from the market weights just established. The equilibrium return on the market reduces from (3.5.22) to a simple

$$\tilde{R}_M(t+1)|_{\vartheta=1} = \bar{\theta}\frac{\tilde{Y}(t+1)}{I(t)} \tag{3.5.34}$$

in case of full depreciation. While the market return in this case does not reflect the same complexities of the capital adjustment technology as (3.5.22), it still reflects the fact that higher future output per unit of current investments implies a higher rate of return. Again this result is consistent with using the findings for the market portfolio weights in the constructing definition of the capital market's return, (3.3.10), and plugging in $\vartheta = 1$. Using (3.5.23) and (3.5.24) the individual securities' equilibrium returns are then

$$\begin{aligned}\tilde{R}_K(t+1)|_{\vartheta=1} &= \left[(1-1)\tilde{\psi}(t+1) + 1 \times \tilde{\theta}(t+1)\right]\frac{\tilde{Y}(t+1)}{I(t)} \\ &= \tilde{\theta}(t+1)\frac{\tilde{Y}(t+1)}{I(t)} \end{aligned} \tag{3.5.35}$$

$$\begin{aligned}\text{and}\quad \tilde{R}_H(t+1)|_{\vartheta=1} &= \left(1-\tilde{\theta}(t+1)\right)\frac{(1-1)\bar{\psi} + 1 \times \bar{\theta}}{1-\bar{\theta}}\frac{\tilde{Y}(t+1)}{I(t)} \\ &= \left(1-\tilde{\theta}(t+1)\right)\frac{\bar{\theta}}{1-\bar{\theta}}\frac{\tilde{Y}(t+1)}{I(t)}. \end{aligned} \tag{3.5.36}$$

As can easily be verified, this is the same as using the market portfolio weights of the degenerate case in the basic formulations of gross returns, (3.2.14) and (3.3.9) respectively. Furthermore, based on (3.5.26) the riskfree rate in the degenerate case of $\vartheta = 1$ is given by

$$\begin{aligned}R(t)|_{\vartheta=1} &= \frac{(1-1)\bar{\psi} + 1 \times \bar{\theta}}{(1-1)\bar{\psi} + 1}\frac{1}{I(t)}\left(\mathbb{E}_t\left[\frac{1}{\left[(1-1)\tilde{\psi}(t+1) + 1\right]\tilde{Y}(t+1)}\right]\right)^{-1} \\ &= \frac{\bar{\theta}}{I(t)}\frac{1}{\mathbb{E}_t\left[\frac{1}{\tilde{Y}(t+1)}\right]}. \end{aligned} \tag{3.5.37}$$

For completeness observe that the stochastic discount factor – for the young or the middle aged – reduces to a handy

$$\begin{aligned}\tilde{M}_y(t+1)|_{\vartheta=1} = \tilde{M}_m(t+1)|_{\vartheta=1} &= \frac{1}{(1-1)\bar{\psi} + \bar{\theta}}\frac{(1-1)\bar{\psi} + 1}{(1-1)\tilde{\psi}(t+1) + 1}\frac{I(t)}{\tilde{Y}(t+1)} \\ &= \frac{I(t)}{\bar{\theta}\tilde{Y}(t+1)}. \end{aligned} \tag{3.5.38}$$

Again this drastically simplified case will be the setting for the numerical example.

3.5.7 Illustrative Example Economy

To continue the simplified example of the illustrative economy, remember from Sect. 3.2.6 that it is characterized by $\vartheta = 1$ and $B = 1$, i.e. full capital depreciation. Thus the above results can be used straightforwardly to calculate equilibrium returns.

Probability Assumptions

However, before doing so, the probabilities for realizing the values specified in Sect. 3.2.6 must finally be specified. Consistent with the independence assumption, let the uncertainty be characterized by the joint probability distribution of Table 3.3.

Table 3.3. Probability distribution in illustrative economy

	$\tilde{\theta}(1) = 0.2$	$\tilde{\theta}(1) = 0.4$	
$\tilde{A}(1) = 10$	$P_1 = 21\%$	$P_3 = 39\%$	$P_{A=10} = 60\%$
$\tilde{A}(1) = 25$	$P_2 = 14\%$	$P_4 = 26\%$	$P_{A=25} = 40\%$
	$P_{\theta=0.2} = 35\%$	$P_{\theta=0.4} = 65\%$	$\sum = 100\%$

Remarks: The table gives the joint probability distribution of the scale and elasticity parameter for the numerical example.

Based on these probabilities the expected value for the elasticity parameter can be calculated as

$$\bar{\theta} = \mathbb{E}_0 \left[\tilde{\theta}(1)\right] = 0.2 \times 35\% + 0.4 \times 65\% = 0.33 \,.$$

Equilibrium Market Weights

Having established $\bar{\theta}$, the market values and the returns follow directly from the results of the previous section and assumptions already made. To start with the market value of the human capital security, recall from Sect. 3.2.6 that investments in $t = 0$ are \$200. Hence the aggregate value of human capital is given by

$$N_m(1)p_H(0) = \frac{1 - 0.33}{0.33} \times \$200 \approx \$406.06 \,.$$

Now observe the internal consistency of the model: in Sect. 3.2.6 it was assumed ad-hoc that the price of a human capital security is \$4.06. This corresponds exactly to the above statement, as solving for $p_H(0)$ with the assumed cohort size[45] of $N_m(1) = N_y(0) = 100$ gives

[45]See Sect. 3.2.6 for thsi assumption.

$$p_H(0) = \frac{\$406.06}{100} \approx \$4.06\,.$$

According to (3.5.33), the market portfolio weights are completely characterized by the expectation for the elasticity parameter:

$$\alpha_K^M(0) = 0.33 \quad \text{and} \quad \alpha_H^M(0) = 1 - 0.33 = 0.67\,.$$

Again, one could derive this result differently. As the value of the physical capital stock is given by the investments, i.e. $K(1)p_K(0) = \$200$, the standard weighting procedure can be used easily. It gives the same results:

$$\alpha_K^M(0) = \frac{K(1)p_K(0)}{K(1)p_K(0) + N_m(1)p_H(0)} = \frac{\$200}{\$200 + \$406.06} = 0.33$$
$$\alpha_H^M(0) = \frac{N_m(1)p_H(0)}{K(1)p_K(0) + N_m(1)p_H(0)} = \frac{\$406.06}{\$200 + \$406.06} = 0.67\,,$$

which again illustrates the consistency on a numerical basis.
Using the formulae from the previous section and the calculations for output $\tilde{Y}(1)$ in the different states one could easily determine each states's return for the real and human capital security. However, this is unnecessary as the calculations in Sects. 3.2.6 and 3.3.6 have already been based on the true values. One must only alter the interpretation, i.e. the possible realizations found there are not *any* returns realizations but those on the human and real capital security *in equilibrium*. Of course the nature of the equilibrium depends on the realizations of $\tilde{A}(1)$ and $\tilde{\theta}(1)$.

Returns and Rates

The return on the capital market can be calculated either by construction from the constituting securities or from the result for the full depreciation case in Sect. 3.5.6. Starting with the latter and using the previous results for output, the gross return on the capital market in the different states is according to (3.5.34)

$$\tilde{R}_M(1) = \begin{cases} \dfrac{0.33 \times \$1,148.70}{\$200.0} \approx \; 190\% & \text{with } P_1 \\ \dfrac{0.33 \times \$2,297.40}{\$200.0} \approx \; 379\% & \text{with } P_2 \\ \dfrac{0.33 \times \$1,319.51}{\$200.0} \approx \; 218\% & \text{with } P_3 \\ \dfrac{0.33 \times \$2,639.02}{\$200.0} \approx \; 435\% & \text{with } P_4\,. \end{cases}$$

In order to illustrate the internal consistency of the approach once again, observe that the above results are identical to those obtained by constructing the market return from its securities. For each case it should hold

$$\tilde{R}_M(1) = \alpha_K^M(0)\tilde{R}_K(1) + \alpha_H^M(0)\tilde{H}_M(1)\,.$$

Taking the numerical results for $\tilde{R}_K(1)$ from Sect. 3.2.6 and for $\tilde{R}_H(1)$ from Sect. 3.3.6 as well as the market weights from above confirms this:

$$\tilde{R}_M(1) = \begin{cases} 0.33 \times 115\% + 0.67 \times 226\% \approx & 190\% & \text{with } P_1 \\ 0.33 \times 230\% + 0.67 \times 453\% \approx & 379\% & \text{with } P_2 \\ 0.33 \times 264\% + 0.67 \times 195\% \approx & 218\% & \text{with } P_3 \\ 0.33 \times 528\% + 0.67 \times 390\% \approx & 435\% & \text{with } P_4\,. \end{cases}$$

Finally, one can calculate the gross riskfree return from (3.5.37) by applying the probabilities defined in Table 3.3. This yields

$$\begin{aligned} R(0) &= \frac{0.33}{\$200}\frac{1}{\frac{1}{\$1{,}148.70}21\% + \frac{1}{\$2{,}297.40}14\% + \frac{1}{\$1{,}319.51}39\% + \frac{1}{\$2{,}639.02}25\%} \\ &= 259\% \end{aligned}$$

which is the numerical value assumed in Sect. 3.4.6. Just for completeness, observe that there is again another way to derive this results. Using the findings for the capital market's returns in (3.5.25) also allows an easy calculation of the riskfree return:

$$\begin{aligned} R(t) &= \frac{1}{\mathbb{E}_1\left[\left(\tilde{R}_M(1)\right)^{-1}\right]} \\ &= \frac{1}{(190\%)^{-1}\,21\% + (379\%)^{-1}\,14\% + (218\%)^{-1}\,39\% + (435\%)^{-1}\,25\%} \\ &= 259\%\,. \end{aligned}$$

Stochastic Discount Factor

Finally, the SDFs for young or middle-aged agents in a model without capital accumulation are given by (3.5.38), i.e.

$$\tilde{M}_y(t+1) = \tilde{M}_m(t+1) = \frac{I(t)}{\theta\tilde{Y}(t+1)}\,.$$

For the illustrative example economy this results in

$$\tilde{M}_y(1) = \tilde{M}_m(1) = \begin{cases} \dfrac{\$200.0}{0.33 \times \$1,148.70} \approx & 0.5276 & \text{with } P_1 \\ \dfrac{\$200.0}{0.33 \times \$2,297.40} \approx & 0.2638 & \text{with } P_2 \\ \dfrac{\$200.0}{0.33 \times \$1,319.51} \approx & 0.4593 & \text{with } P_3 \\ \dfrac{\$200.0}{0.33 \times \$2,639.02} \approx & 0.2297 & \text{with } P_4\,. \end{cases}$$

Checking with these values for instance the relation (3.5.28) for human capital, which is the SDF-representation of the numerical illustrations in Sect. 3.4.6,

$$\begin{aligned}\mathbb{E}_0\left[\tilde{M}_y(1)\tilde{R}_H(1)\right] = \; & 0.5276 \times 226\% \times 21\% + 0.2638 \times 453\% \times 14\% \\ & +0.4593 \times 195\% \times 39\% + 0.2297 \times 390\% \times 26\% \approx 1,\end{aligned}$$

is just another illustration for the correctness of the approach.

3.6 General Equilibrium

The equilibrium conditions derived so far are not yet proper reduced-form equations for the equilibrium returns, as they still depend on endogenous variables, like output. In order to derive such conditions it is necessary to establish equilibrium in the economy's other markets as well.

Due to the labor market' simplicity, its equilibrium is trivial and has already been established in Sect. 3.3.2. With only consumption but not leisure entering utility in (3.4.4), the middle-aged cohort dedicates its entire time to working so that this cohort's size determines the fixed labor supply. Based on the standard neoclassical assumption of competitiveness, adjustments of the wage rate allow the market to clear. Hence, in formal language labor market equilibrium is characterized by

$$L(t) = L^D(t) = L^S(t) = N_m(t),$$

which is a restatement of (3.3.6).

Because the model abstracts from international and monetary aspects neither the foreign exchange nor the money market need to be cleared. Equilibrium on the capital market has just been established so that the remaining and decisive aspect of the economy is the goods market. In this market, both the consumptive and capital accumulating dimensions of the consumers choice are reflected. The remainder of this section is dedicated to establishing equilibrium in it and showing how real and financial decisions translate consistently into general equilibrium.

The demand for goods is driven by consumption and investments in physical capital. The first section analyzes the former, the second will do it on the latter and the subsequent section finally uses the preliminary results to derive a proper reduced form expression for output.

3.6.1 Consumption by Cohorts

All three cohorts consume goods. Based on earlier findings the generational consumption aggregates can be related to investments or output.

Consumption of the Young

The young cohort's aggregate consumption has been established in (3.5.3). Now plug in this the definition of aggregate wealth of the young, (3.5.1), and remember that a young agent's initial wealth endowment is given by his human capital, (3.4.8). Making additionally usage of the certain demographic structure implies that consumption of the young can be written as

$$C_y(t) = \frac{1}{\beta^2+\beta+1} W_y(t) = \frac{1}{\beta^2+\beta+1} N_m(t+1) p_H(t) \,.$$

This can be expressed as

$$= \frac{1-\bar{\theta}}{(\beta^2+\beta+1)} \frac{I(t)}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right]} \,, \tag{3.6.1}$$

where the human capital's market capitalization in equilibrium, (3.5.17), has been used to relate this cohort's consumption and aggregate investments in real capital.

Consumption of the Middle-aged

One could try similar steps for the consumption of the middle-aged cohort, but this requires knowledge of their current wealth. Yet, in order to specify this $W_m(t)$ one would need to characterize the entire history of the economy. This way one could establish a founded representation for the realized returns $R_H(t)$ and $R_K(t)$, the portfolio weights $\alpha_{H,y}(t-1)$ and $\alpha_{K,y}(t-1)$ and the middle-aged previous savings $s_y(t-1)$. In order to avoid this, an elegant way is using aggregate savings and establishing the middle-aged's contribution to it backwards.
Therefore, observe first that aggregate savings must equal the market value of all securities. The reason for this is rather simple: in equilibrium all securities must be held at their equilibrium prices. As financial instruments are purchased only for saving motives it must hold that

$$S(t) = K(t+1)p_K(t) + N_m(t+1)p_H(t) \,. \tag{3.6.2}$$

Using the finding for the total market value, (3.5.19), allows to rewrite this as

$$S(t) = \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\vartheta\bar{\psi}+\vartheta^2\bar{\theta}} I(t) \,. \tag{3.6.3}$$

Section 5.2.2 will give further insights how this compares to the standard economic notion that savings and investments *equal.* In order to derive a statement for $C_m(t)$, split the left hand side of (3.6.3) into the contributions from the young and from the middle-aged by using definition (3.5.2). Then substitute the young generation's savings with $W_y(t) - C_y(t)$ according to (3.5.5) and solve for $S_m(t)$. Replacing the consumption of the young by (3.6.1) and their wealth by $(1-\bar{\theta})/[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}]I(t)$ for the same reasoning as above yields

$$\begin{aligned} S_m(t) &= S(t) - W_y(t) + C_y(t) \\ &= \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\vartheta\bar{\psi}+\vartheta^2\bar{\theta}} I(t) - \frac{1-\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} I(t) \\ &\quad + \frac{1}{\beta^2+\beta+1} \frac{1-\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} I(t). \end{aligned} \tag{3.6.4}$$

As shown in App. B.3.1, this can be simplified to

$$S_m(t) = \left[\frac{1}{\vartheta} + \frac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)}\right] I(t)\,. \tag{3.6.5}$$

Because $C_m(t) = 1/(\beta+1)W_m(t)$ and $S_m(t) = \beta/(\beta+1)W_m(t)$ according to (3.5.3) and (3.5.4), it follows that

$$C_m(t) = \frac{1}{\beta} S_m(t) = \left[\frac{1}{\vartheta} + \frac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)}\right] \frac{I(t)}{\beta}\,. \tag{3.6.6}$$

This relates consumption of the middle-aged generation to investments.

Consumption of the Old

Finally, an expression for $C_o(t)$ can be derived. Take the old's aggregated policy, (3.5.3), and replace $W_o(t) = N_o(t)w_o(t)$ by the optimal portfolio's return on previous savings. In (3.5.13) it has been found that this return is identical to the market return. Replacing $s_m(t-1)N_m(t-1)$ by $S_m(t-1)$ allows to apply (3.6.5):

$$\begin{aligned} C_o(t) &= W_o(t) = N_o(t)w_o(t) = N_o(t)s_m(t-1)R^*_{P,m}(t) = S_m(t-1)R_M(t) \\ &= \left[\frac{1}{\vartheta} + \frac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)}\right] I(t-1)R_M(t)\,. \end{aligned}$$

Plugging in period t's realized market return from (3.5.22), one can establish

$$\begin{aligned} C_o(t) &= \left[\frac{1}{\vartheta} + \frac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)}\right] I(t-1) \\ &\quad \times \left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right] \frac{(1-\vartheta)\psi(t)+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta} \frac{Y(t)}{I(t-1)} \\ &= \left[\frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{\vartheta} + \frac{1-\bar{\theta}}{\beta^2+\beta+1}\right] \frac{(1-\vartheta)\psi(t)+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta} Y(t)\,. \end{aligned} \tag{3.6.7}$$

One could now try to sum across all generations to establish the economy's aggregate consumption. However, while $C_o(t)$ depends on output, $C_m(t)$ and $C_y(t)$ are functions of investments. The next section establishes a relation between these two variables.

3.6.2 Investment-Output Ratio

With relations connecting each generation's consumption to output or investment, one can apply them in the national income identity.

Identifying $\psi(t)$

Starting with (3.2.3), $C(t)$ can be split according to (3.5.2). Then (3.6.1), (3.6.6) and (3.6.7) are substituted in for $C_y(t)$, $C_m(t)$ and $C_o(t)$ respectively:

$$\begin{aligned} Y(t) &= C(t) + I(t) = C_y(t) + C_m(t) + C_o(t) + I(t) \\ &= \frac{1-\bar{\theta}}{(\beta^2+\beta+1)\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right]} I(t) \\ &+ \left[\frac{1}{\vartheta} + \frac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)}\right] \frac{I(t)}{\beta} \\ &+ \left[\frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{\vartheta} + \frac{1-\bar{\theta}}{\beta^2+\beta+1}\right] \frac{(1-\vartheta)\psi(t)+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta} Y(t) + I(t). \end{aligned} \quad (3.6.8)$$

Bringing all $Y(t)$-terms to the left side and dividing by $Y(t)$ results in

$$\begin{aligned} 1 - &\left[\frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{\vartheta} + \frac{1-\bar{\theta}}{\beta^2+\beta+1}\right] \frac{(1-\vartheta)\psi(t)+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta} \\ &= \left[\begin{array}{c} \dfrac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)} + \dfrac{1}{\beta\vartheta} \\ + \dfrac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)\beta} + 1 \end{array}\right] \psi(t). \end{aligned} \quad (3.6.9)$$

where $\psi(t) = I(t)/Y(t)$ from Sect. 3.5.3 has been used. Equation (3.6.9) can be solved for $\psi(t)$. As App. B.3.2 shows, the solution is

$$\psi(t) = \left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right] \frac{(1-\bar{\theta})\beta^2\vartheta^2}{\Xi_1}, \quad (3.6.10)$$

where Ξ_1 is a constant given by

$$\Xi_1 \equiv \left\{\begin{array}{l} \left[(1-\vartheta)\bar{\psi}+\vartheta\right]^2 \\ + \left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right]\beta\left[(1-\vartheta)\bar{\psi}+\vartheta\right] \\ + \left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right]^2\beta^2 \\ + \left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](1-\bar{\theta})\beta^2\vartheta^2 \end{array}\right\}. \quad (3.6.11)$$

But now note: as (3.6.10) has *only* parameters on the right hand side, the investment-output ratio is truly independent of t, i.e. $\psi(t) = \psi$. This gives a final justification of the assumptions made in Sect. 3.5.3. There it was assumed that agents decide based on the expectation of an *effectively* constant mean of the investment-output ratio. But if the investment-output ratio is a constant as just shown, this assumption becomes trivial because

$$\mathbb{E}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\bar{\psi}}\right]=\mathbb{E}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\psi}\right]=\frac{1}{1+\frac{1-\vartheta}{\vartheta}\psi} \tag{3.6.12}$$

and combing this with (3.5.16) implies

$$\frac{1}{1+\frac{1-\vartheta}{\vartheta}\bar{\psi}}=\frac{1}{1+\frac{1-\vartheta}{\vartheta}\psi} \quad \text{and thus} \quad \bar{\psi}=\psi. \tag{3.6.13}$$

The actual investment-output ratio and the assumed $\bar{\psi}$ must be identical![46] In other words, one could have guessed $\psi(t)=\psi$ first and then proofed this by deriving (3.6.10). Instead of solving the explicit form of (3.6.10), i.e.

$$\psi=\frac{1}{\frac{[(1-\vartheta)\psi+\vartheta]^2}{\left[(1-\vartheta)\psi+\vartheta\bar{\theta}\right](1-\bar{\theta})\beta^2\vartheta^2}+\frac{[(1-\vartheta)\psi+\vartheta]}{(1-\bar{\theta})\beta\vartheta^2}+\frac{\left[(1-\vartheta)\psi+\vartheta\bar{\theta}\right]}{(1-\bar{\theta})\vartheta^2}+1}, \tag{3.6.14}$$

which is cumbersome and unnecessary, let the relevant solution of this equation be

$$\psi=\bar{\psi}=\Psi\left(\beta,\vartheta,\bar{\theta}\right)\doteq\Psi\,. \tag{3.6.15}$$

This is a noteworthy result: the fraction of output that agents do not consume but assign to investments in real capital is determined by the subjective time preference, the form of the elasticity in the capital adjustment technology and the expected elasticity in the consumption technology. Remembering furthermore that for the assumed logarithmic utility the coefficient of relative risk aversion is fixed at unity, one notices that (3.6.15) is governed by those parameters which describe the utility and the production function. This reflects the intuition of Fig. 2.2: on the one hand, utility maximizing agents optimize their intertemporal consumption allocation; based on these implicit investment decisions firms can, on the other hand, transform present consumable units into future output. Ψ reflects the result of both of these actions.[47]

[46]Naturally the covariance of a constant with a random variable is zero, i.e.

$$\mathrm{Cov}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\bar{\psi}},\tilde{\theta}\right]=\mathrm{Cov}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\psi},\tilde{\theta}\right]=0\,,$$

which is the second part of (3.5.16).

[47]Clearly, the typical critique can be applied here: only if agents act based on the presented model it is rational to assume the long-run property of $\bar{\psi}$ as assumed. Nevertheless – as for many neoclassical models – this should not be a reason to dismiss the approach per se.

Properties of ψ

Given the boundaries of β, ϑ and $\bar{\theta}$ the possible values for Ψ are limited and can easily be summarized graphically. Figure 3.3 gives the values of the investment-output ratio for $\beta = 0.6$. Appendix B.3.2 has illustrations for alternative values of β. Due to the generality of the technology for capital accumulation, the figure also nests the special cases of physical capital's full depreciation and the Lucas [1978] exchange economy.

As can be seen, Ψ is smaller than one for all parameter combinations. This is satisfying since by construction investment in physical capital must not exceed output. Still, even with other values for β, the maximum values of the investment-output ratio are only around 10%. This is rather low compared to typical observed numbers of about 30%. A reason for this is the period length. With a time step of 30 years only few expenditures qualify as investments. One can reasonably assume that adding more generations to the model and hence reducing the period length would push Ψ into the right direction.[48]

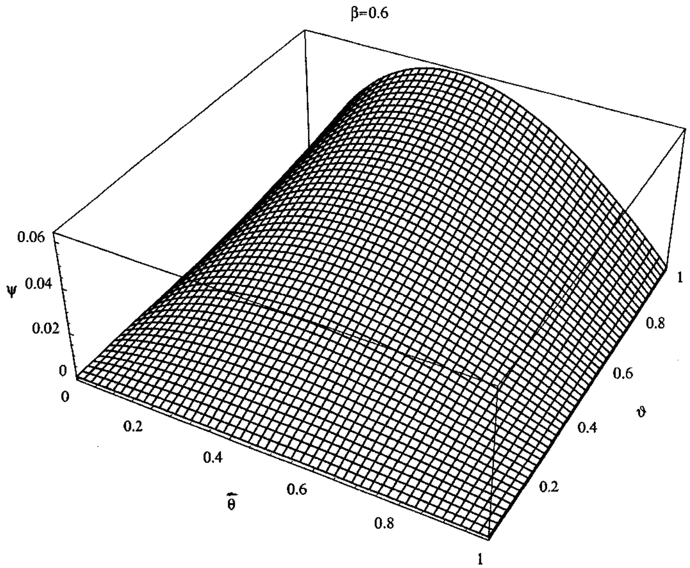

Fig. 3.3. Numerical values of investment-output ratio

Remarks: The plot visualizes the values of the investment-output ratio Ψ for different values of the technologies' elasticity parameters ϑ and $\bar{\theta}$ and a subjective discount factor of $\beta = 0.6$. It also embeds the Lucas [1978] fruit-tree model at $\vartheta = 0$ and the case of full depreciation at $\vartheta = 1$.

[48]The intuition for this stems from the consumption policies, (3.4.19) and (3.4.28). With shorter periods, β will increase making the fraction of wealth consumed smaller. Additionally, the number of generations would increase, resulting in higher

Limits of ψ

As the investment-output ratio integrates a rather general setting it has interesting limits. Starting with the limiting values with respect to the consumption technology, Fig. 3.3 illustrates that the investment-output ratio equals zero if the expected elasticity parameter in the consumption technology is either zero or unity.[49]

$$\lim_{\bar{\theta}\to 0} \Psi = 0 \qquad \lim_{\bar{\theta}\to 1} \Psi = 0 \tag{3.6.16}$$

In the case of $\bar{\theta} = 0$ physical capital is not expected to contribute to production of consumable units in (3.2.1); i.e. the macroeconomic production function degenerates approximately to $Y(t) = \tilde{A}(t)L(t)$.[50] With human labor being the only expected input factor, physical capital is not expected to earn any rental and is thus considered worthless. Hence it is senseless for investors to allocate savings to real capital and no investment in physical capital will be realized.

In the other extreme of $\bar{\theta} = 1$ the investment to output ratio approaches zero as well. In this case human labor is not utilized in the consumption technology as (3.2.1) becomes roughly $Y(t) = \tilde{A}(t)K(t)$. But with expectedly worthless labor agents neither expect to earn a labor income when middle-aged nor can they capitalize on it when young. This means that their wealth is zero and hence they cannot save anything, which implies that there are no investments in real capital either.

Turning to the extreme cases of the capital adjustment technology the *exchange*-character of the Lucas [1978] fruit-tree model is evident as

$$\lim_{\vartheta\to 0} \Psi = 0. \tag{3.6.17}$$

With $\vartheta = 0$ the physical capital stock becomes independent of investments and remains constant over time as shown in (3.2.9). As such a tree bears fruits regardless of investments, agents cannot increase their utility by saving and investing in real capital which implies $\Psi = 0$. This emphasizes the exchange character of the Lucas [1978] setting: without the alternative of saving agents can only exchange their endowments.

For the degenerate case of full real capital depreciation one obtains as shown again in App. B.3.2

$$\Psi|_{\vartheta=1} = \frac{\beta^2\bar{\theta}(1-\bar{\theta})}{1+\beta\bar{\theta}+\beta^2\bar{\theta}}\,. \tag{3.6.18}$$

---

absolute savings, which is not captures by the the model and hence by the illustrations of various βs.

[49]See also App. B.3.2.

[50]This is only an explanatory interpretation: in (3.2.1) actually the realization of $\tilde{\theta}$ enters and this could be distinct from $\bar{\theta}$. Thus, the macroeconomic production function degenerates only in an expected sense regarding the elasticity parameter.

Finally, observe the effect of infinity impatience by agents. In this case they derive utility only from *present* consumption. In the utility model of Sect. 3.4.1 this implies that the subjective time preference parameter β goes to zero, making future consumption worthless in today's utils. In this case, the limit of the investment-output ratio is

$$\lim_{\beta \to 0} \Psi = 0. \tag{3.6.19}$$

This is an intuitive result: when agents are infinitely impatient they immediately consume their entire wealth, which translates into zero savings. Hence, there are no investments which implies an investment-output ratio of zero as long as un-depreciated physical capital allows to produce.[51]

3.6.3 Goods Market Equilibrium

Finally, the solution of the investment-output ratio allows to establish true reduced-form expressions for output and gross domestic product. These imply equilibrium on the market for consumption goods.

Output

Based on the solved investment-output ratio of (3.6.15) investment in physical capital is characterized by

$$I(t) = \psi(t)Y(t) = \Psi Y(t). \tag{3.6.20}$$

This allows to connect the real capital stock and macroeconomic output by inserting (3.6.20) in the capital adjustment technology from Sect. 3.2.2. Equation (3.2.7) becomes then

$$K(t+1) = B\Psi^{\vartheta} Y(t)^{\vartheta} K(t)^{1-\vartheta}. \tag{3.6.21}$$

As this can also be formulated for the present stock $K(t)$, (3.6.21) allows to solve for the physical capital stock based on initial real capital K_0 and past output; i.e. $K(t)$ is given by a function $\check{K}$

$$K(t) \equiv \check{K}(\Psi, B, K_0, Y, t-1) \doteq \check{K}(t-1). \tag{3.6.22}$$

Using this and the labor market equilibrium of (3.3.6) in the assumed consumption technology (3.2.1), period t's output can be characterized as

$$Y(t) = A(t)\check{K}(t-1)^{\theta(t)} N_m(t)^{1-\theta(t)} \equiv \Upsilon\left(\theta, A, \check{K}, N_m, t\right) \doteq \check{Y}(t), \tag{3.6.23}$$

where $\check{Y}(t)$ is a function whose arguments are all known at time t. In other words, once the uncertainties of period t, namely $\tilde{A}(t)$ and $\tilde{\theta}(t)$, have been resolved, output $Y(t)$ is completely determined by the function $\check{Y}(t)$.

[51]Due to the approximate specification of (3.2.7) this is only possible in a single period. With an actual additive relationship between net investment and depreciation, the stock of physical capital would decay exponentially and allow so for future consumption, too, even with infinity impatience.

Reduced Form Representations

Principally, one could run over all findings and replace $Y(t)$ by $\check{Y}(t)$ – and $\bar{\psi}$ as well as $\psi(t+1)$ by Ψ according to (3.6.15) – in order to derive truly complete solutions for all endogenous variables. For instance, one could reexpress (3.6.22) by using the functional form for output. Fortunately, this is generally not necessary. Still, some variables are of interest for an explicit solution, like the presentation of the gross domestic product. In Sect. 3.3.1, $Y^{GDP}(t)$ was found to be given by

$$Y^{GDP}(t) = Y(t) + \frac{1-\vartheta}{\vartheta} I(t).$$

Applying the investment-output ratio and functional representation for output, it follows that

$$Y^{GDP}(t) = Y(t) + \frac{1-\vartheta}{\vartheta} \Psi Y(t) = \left(1 + \frac{1-\vartheta}{\vartheta}\Psi\right) \Upsilon(t) \,. \tag{3.6.24}$$

Ignoring the limiting cases, gross domestic product is always higher than output since in general $\Psi > 0$ and $0 < \vartheta < 1$. As mentioned earlier this is the consequence of the convex costs in the capital adjustment technology.
For the further analysis in Chap. 4 three more explicit reduced form equations are required and can easily be derived here. From (3.6.1), (3.6.6) and (3.6.7) consumption of the young, middle-aged and old generation can be restated using the investment-output ratio and (3.6.23) as

$$C_y(t) = \frac{(1-\bar{\theta})\Psi}{(\beta^2+\beta+1)\left[(1-\vartheta)\Psi + \vartheta\bar{\theta}\right]} \check{Y}(t), \tag{3.6.25}$$

$$C_m(t) = \left[\frac{1}{\vartheta} + \frac{1-\bar{\theta}}{(\beta^2+\beta+1)\left[(1-\vartheta)\Psi + \vartheta\bar{\theta}\right]}\right] \frac{\Psi}{\beta} \check{Y}(t), \tag{3.6.26}$$

and

$$C_o(t) = \left[\frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{\vartheta} + \frac{1-\bar{\theta}}{\beta^2+\beta+1}\right] \check{Y}(t). \tag{3.6.27}$$

Equilibrium

Finally, note that equilibrium in the goods market has been established rather implicitly. The investment-output ratio is derived in such a way that it automatically clears the goods market. This follows from the fact that in Sect. 3.6.2 the introduction of Ψ is essentially based on applying the national income identity. But this identity is nothing else than the statement that supply and demand for goods equal. The supply is given by the period's production of consumable units, $Y(t)$, and demand consists of aggregated consumption, $C(t)$, and dedication of consumption units into the accumulation of physical capital via investments, $I(t)$. Hence, the solution of Ψ is a representation of

equilibrium in the goods market and this variable will play key a role when replicating the complete market results with PAYGO in Chap. 4.
A comprehensive summary of all equilibrium findings will be given in Sect. 3.7.

3.6.4 Full Depreciation

Once again, the case of a fully depreciating physical capital stock shall be examined in more detail.

Consumption, Saving and Investment

According to (3.6.1), (3.6.6) and (3.6.7), the degenerate capital adjustment function implies that consumption of the three generations is given by

$$\begin{aligned} C_y(t)|_{\vartheta=1} &= \frac{1-\bar{\theta}}{(\beta^2+\beta+1)} \frac{I(t)}{\left[(1-1)\bar{\psi}+1\times\bar{\theta}\right]} \\ &= \frac{1-\bar{\theta}}{(\beta^2+\beta+1)\bar{\theta}} I(t), \qquad (3.6.28) \\ C_m(t)|_{\vartheta=1} &= \left[\frac{1}{1} + \frac{1-\bar{\theta}}{\left[(1-1)\bar{\psi}+1\times\bar{\theta}\right](\beta^2+\beta+1)}\right] \frac{I(t)}{\beta} \\ &= \frac{\beta(\beta+1)\bar{\theta}+1}{(\beta^2+\beta+1)\beta\bar{\theta}} I(t), \qquad (3.6.29) \\ \text{and} \quad C_o(t)|_{\vartheta=1} &= \left[\frac{(1-1)\bar{\psi}+1\times\bar{\theta}}{1} + \frac{1-\bar{\theta}}{\beta^2+\beta+1}\right] \frac{(1-1)\psi(t)+1}{(1-1)\bar{\psi}+1} Y(t) \\ &= \frac{\beta(\beta+1)\bar{\theta}+1}{\beta^2+\beta+1} Y(t). \qquad (3.6.30) \end{aligned}$$

Plugging these equations in the condition of equilibrium in the goods market yields

$$\begin{aligned} Y(t) &= C_y(t)|_{\vartheta=1} + C_m(t)|_{\vartheta=1} + C_o(t)|_{\vartheta=1} + I(t) \\ &= \frac{1-\bar{\theta}}{(\beta^2+\beta+1)\bar{\theta}} I(t) + \frac{\beta(\beta+1)\bar{\theta}+1}{(\beta^2+\beta+1)\beta\bar{\theta}} I(t) + \frac{\beta(\beta+1)\bar{\theta}+1}{\beta^2+\beta+1} Y(t) + I(t). \end{aligned} \qquad (3.6.31)$$

As the degenerate case is a very different economy, it is more accurate to use $Y(t)|_{\vartheta=1}$ and $I(t)|_{\vartheta=1}$. Applying again the investment-output ratio $\psi(t)|_{\vartheta=1} = I(t)|_{\vartheta=1}/Y(t)|_{\vartheta=1}$ allows to restate (3.6.31) as

$$\begin{aligned} 1 &= \frac{1-\bar{\theta}}{(\beta^2+\beta+1)\bar{\theta}} \psi(t)|_{\vartheta=1} + \frac{\beta(\beta+1)\bar{\theta}+1}{(\beta^2+\beta+1)\beta\bar{\theta}} \psi(t)|_{\vartheta=1} \\ &\quad + \frac{\beta(\beta+1)\bar{\theta}+1}{\beta^2+\beta+1} + \psi(t)|_{\vartheta=1} \,. \end{aligned} \qquad (3.6.32)$$

According to App. B.3.3, the solution of the above equation is given by

$$\psi(t)|_{\vartheta=1} = \frac{\beta^2\bar{\theta}(1-\bar{\theta})}{1+\beta\bar{\theta}+\beta^2\bar{\theta}}\,. \tag{3.6.33}$$

Again, this is independent of time and comparing it with the general case's limit for $\vartheta = 1$ in (3.6.18) shows the consistency of the results because

$$\psi(t)|_{\vartheta=1} = \psi|_{\vartheta=1} = \Psi|_{\vartheta=1}\,. \tag{3.6.34}$$

Output and Gross Domestic Product

Since the capital adjustment technology in the degenerate case is not only characterized by $\vartheta = 1$, but also requires $B = 1$,[52] the intertemporal connection of physical capital stock and output reduces from (3.6.22) to

$$\begin{aligned} K(t)|_{\vartheta=1} &= \Psi|_{\vartheta=1} Y(t-1)|_{\vartheta=1} \\ &= \breve{K}|_{\vartheta=1}\left(\Psi, \vartheta=1, Y, t-1\right) \doteq \breve{K}(t-1)|_{\vartheta=1}\,. \end{aligned} \tag{3.6.35}$$

Note the difference to the general case: with full depreciation of physical capital the initial capital stock does not enter in (3.6.35).
Using the last result in the consumption technology (3.2.1), which is affected by the special case only via the real capital stock, allows to establish the full reduced form representation as

$$\begin{aligned} Y(t)|_{\vartheta=1} &= A(t)\breve{K}|_{\vartheta=1}(t-1)^{\theta(t)} N_m(t)^{1-\theta(t)} \\ &\equiv \Upsilon\left(\theta, \vartheta=1, A, \breve{K}, N_m, t\right) \doteq \breve{Y}(t)|_{\vartheta=1}\,. \end{aligned} \tag{3.6.36}$$

One could make the exercise and rewrite everything with $\breve{Y}(t)|_{\vartheta=1}$ instead of $Y(t)|_{\vartheta=1}$ again. Noteworthy is only that gross domestic product and output are equal in the degenerate case, i.e. $Y^{GDP}(t)|_{\vartheta=1} = \breve{Y}(t)|_{\vartheta=1}$. This was already shown in Sect. 3.3.5 but is also implied by reduced form representation of (3.6.24).

3.6.5 Illustrative Example Economy

Having reduced the goods' market equilibrium of the general model to the case without capital accumulation, the numerical example of the illustrative economy can be continued from the preceding sections. Therefore, extend it by the following additional assumptions.
The example economy's population is characterized by a declining birth rate so that its present demographic structure is given by $N_y(0) = 100$ young agents, $N_m(0) = 125$ middle-aged agents and $N_o(0) = 156$ old agents. The young generation's size corresponds to the assumed labor force for $t = 1$

[52] See Sect. 3.2.2 and remember that $B = 1$ has been – and will be – dropped from the notation only for convenience.

of Sect. 3.2.6. Note that with the 30-year time horizon these demographics correspond to a rate of decline in yearly cohort sizes of only about 0.75%! Based on these assumptions, the results of Sect. 3.4.6 can be aggregated to cohort variables according to (3.5.1):

$$\begin{aligned} C_y(0) &= \$2.07 \times 100 \approx \$207.17 & S_y(0) &= \$4.06 \times 100 \approx \$198.89 \\ C_m(0) &= \$5.43 \times 125 \approx \$678.62 & S_m(0) &= \$3.26 \times 125 \approx \$407.17 \\ C_o(0) &= \$14.25 \times 156 \approx \$2,222.91 & S_o(0) &= \$0.00 \times 156 \approx \$0.00 \end{aligned}$$

Summing these values to the economy's aggregated variables as defined in (3.5.2) yields that total consumption in $t = 0$ is

$$C(0) = \$207.17 + \$678.62 + \$2,222.91 \approx \$3.108,71 .$$

According to the national income identity, (3.2.3), output in $t = 0$ is thus given by

$$Y(0) = C(0) + I(0) = \$3.108,71 + \$200.00 = \$3,308.71 ,$$

where $I(0) = \$200.00$ as assumed in Sect. 3.2.6. Hence, the economy's *constant* investment-output ratio can be calculated as

$$\Psi = \frac{\$200.00}{\$3,308.71} \approx 6.04\% .$$

Now observe that this is the same as using the degenerate case's explicit result for the investment-output ratio. That is, applying the assumed parameter value $\beta = 0.6$ and the expected elasticity of $\bar{\theta} = 0.33$ from Sect. 3.5.7 in (3.6.18) gives again

$$\Psi = \frac{0.6^2 \times 0.33 \times (1 - 0.33)}{1 + 0.6 \times 0.33 + 0.6^2 \times 0.33} \approx 6.04\% .$$

Chapter 5 will elaborate on this investment output ratio to further illustrate the model's internal consistency and connect it to the replication.

3.7 Summary for Tradable Human Capital

In this chapter, a general equilibrium model with overlapping generations and tradable human capital has been developed. The description has been based on a sequential order of setting the framework, deriving relationships between macroeconomic and financial variables, analyzing the optimization rationales, aggregating the consequences and establishing equilibrium. This section briefly summarizes the major findings and restates important formal results in Tab. 3.4.

Table 3.4. Summary for tradable human capital

	Equilibrium Conditions		
Goods Market	$Y(t)$	$= C(t) + I(t) = \check{Y}(t)$	(3.2.3)
Investments	$I(t)$	$= \Psi \times \check{Y}(t)$	(3.6.15)
Capital Market	$\alpha_K^M(t)$	$= \frac{(1-\vartheta)\Psi+\vartheta\bar{\theta}}{(1-\vartheta)\Psi+\vartheta}$	(3.5.20)
	$\alpha_H^M(t)$	$= \frac{(1-\bar{\theta})\vartheta}{(1-\vartheta)\Psi+\vartheta}$	(3.5.21)
	Optimization by Agents		
Allocations	$\alpha_{j,a}^*(t)$	$= \alpha_j^M(t) \quad \text{for } a = y, m, o;\ j = K, H$	(3.5.6)
Savings	$S_y(t)$	$= \frac{\beta^2+\beta}{\beta^2+\beta+1} W_y(t),$	(3.5.4)
	$S_m(t)$	$= \frac{\beta}{\beta+1} W_m(t)$ and $S_o(t) = 0$	
Consumption	$C_y(t)$	$= \frac{1}{\beta^2+\beta+1} W_y(t),$	(3.5.3)
	$C_m(t)$	$= \frac{1}{\beta+1} W_m(t)$ and $C_o(t) = W_o(t)$	
Wealth Budget	$W_a(t)$	$= C_a(t) + S_a(t) \quad \text{for } a = y, m, o$	(3.5.5)
	Returns in Financial Markets		
Riskless Asset	$R(t)$	$= \frac{\frac{(1-\vartheta)\Psi+\vartheta\tilde{\theta}}{[(1-\vartheta)\Psi+\vartheta]\Psi Y(t)}}{\mathbb{E}_t\left[\frac{1}{[(1-\vartheta)\Psi+\vartheta]\tilde{Y}(t+1)}\right]}$	(3.5.26)
Real Capital	$\tilde{R}_K(t+1)$	$= [(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)]\frac{\tilde{Y}(t+1)}{\Psi Y(t)}$	(3.5.23)
Human Capital	$\tilde{R}_H(t+1)$	$= [(1-\vartheta)\Psi + \vartheta\bar{\theta}]\frac{1-\tilde{\theta}(t+1)}{1-\bar{\theta}}\frac{\tilde{Y}(t+1)}{\Psi Y(t)}$	(3.5.24)
Capital Market	$\tilde{R}_M(t+1)$	$= [(1-\vartheta)\Psi + \vartheta\bar{\theta}]\frac{\tilde{Y}(t+1)}{\Psi Y(t)}$	(3.5.22)
	Macroeconomic Framework		
Gross Domestic Product	$Y^{GDP}(t)$	$= C(t) + \frac{1}{\vartheta} I(t)$	(3.3.4)
Consumable Output	$Y(t)$	$= \tilde{A}(t) \times K(t)^{\tilde{\theta}(t)} \times L(t)^{1-\tilde{\theta}(t)}$	(3.2.1)
Capital Accumulation	$K(t+1)$	$= B \times I(t)^{\vartheta} \times K(t)^{1-\vartheta}$	(3.2.7)
Generational Change	$N_m(t+1)$	$= N_y(t)$	(3.1.2)

Remarks: The table summarizes important results of the model with tradable human capital and gives references to the respective equations.

Starting at the table's bottom, the model is based on a macroeconomic framework: this determines the economy's stochastic intertemporal terms-of-trade by means of physical capital accumulation and demographic changes. Taking a specific amount of physical capital and a certain labor force as resources into a new period enables the production of consumable units according to the random specifications of the consumption output technology. Preference-based valuation of this output allows a statement of the gross domestic product.
This macroeconomic framework translates into return expectations and thus prices in the economy's financial markets. Due to the assumption of completeness, there are assets on both devices of intertemporal resource transfer: a human capital security and a real capital security. While the intertemporal terms-of-trade determine the general performance of the capital market, its cross-sectional distribution is influenced by each factor's contribution, fluctuating randomly. This stochastic component allows also to establish the notion of a riskfree rate representing the shadow price of a nonexistent certainty equivalent device of intertemporal resource transfer.
Returns on the financial markets, realized as well as expected ones, determine the distribution of aggregated wealth between the young, middle-aged and old generation. The resulting wealth budgets are the constraints for the intertemporal utility maximization of each cohorts' members. This optimization yields the desired levels of current consumption and savings as well as the allocation of the savings among the different securities for the purpose of diversification.
To achieve equilibrium in all markets the agents' aggregated decisions must match the terms-of-trade offered by the economy's macroeconomic framework and reflected in its financial markets. On the capital market, agents must be willing to hold the aggregated value of all securities. The agents' optimization – unconstrained by market incompleteness – results in identical allocations to both securities, which thus reflect the weights on the capital market. As this translates into an amount of investment that corresponds to the return expectations on the different securities, the consequences of the intertemporal optimization are kicked back into the economy's macroeconomic framework, where consumption and physical investment decisions must clear the market of consumable output.
In other words, the model established in this chapter allows a consistent specification of the intertemporal development of an economy's consumption and production side under uncertainty with overlapping generations.

4

Replication with PAYGO

Raw demography is creating promises of implicit assets faster than the underlying economy is growing.

Kotlikoff and Burns [2004]: "The Coming Generational Storm", p. 204

Having established the benchmark with tradable human capital in Chap. 3, this chapter shows how this solution can be replicated when there is no security on human capital. By introduction of a pay-as-you-go pension system and a consumption tax to benefit the young, exactly the same allocations of consumption and investment are achievable as in the benchmark: the young can consume and invest in physical capital and the old participate in the factor income of labor. While the standard approaches of solving this replication on an individual basis fail, the aggregation to the macroeconomic level allows to establish the necessary calibration of the tax rate and of the contributions to the pension system.

4.1 Incomplete Markets and Overview

In order to fully appreciate the welfare enhancing effect of the PAYGO introduction, it is crucial to understand the artificial nature of the framework presented in Chap. 3. The security tracking the value of human capital is purely hypothetical; it does not exist in reality. Hence, the first-best solution to the old-age savings problem proposed in the last chapter is not achievable. However, as discussed in Sect. 2.3 the idea of this analysis is the replication of the complete markets' solution in a second-best approach. This section reflects further on this argument and sketches an overview of the model for replicating human capital's tradability with a pay-as-you-go pension system.

4.1.1 First- versus Second-best

The benchmark model's solution rests heavily on the assumption of a security perfectly tracking the value of human capital. The hypothetical existence of a liquid market for these securities allows all agents to trade claims on labor's share of future output. Sects. 3.4.3 and 3.5.4 have shown that these claims are advantageous for all agents who engage in saving. As the human capital securities follow stochastics that are different from those of physical capital they are a mean for diversification. Hence, it is rational for the young as well as for the middle-aged to form portfolios consisting of both securities. The hypothetical human capital securities also enables the analysis of the intertemporal consumption-savings problem within the standard framework of finance. Instead of adding income from labor to the appreciated value of the savings portfolio, one can include it in the portfolio consideration. Hence, from a technical point of view the human capital securities prevent the troublesome non-additivity problem shown in (2.2.4). Finally, they are also the only mean allowing the young generation to consume. Because a young agent does not work and families play no role in this framework, issuance of human capital securities is the only source of funds for him. From the proceeds he is not only able to acquire ownership of physical capital but also – and more importantly – to pay for his consumption. An intuitive interpretation of tradability is thus that young agents borrow against the *entire* value of their human capital and reinvest only the desired amount of savings in it.[1]

Nontradable Human Capital

However, in reality a young person is prevented from doing so because even in modern economies human capital alone is not considered as sufficient collateral. The reasons for this market failure have been indicated in Sect. 2.2.2: moral hazard and adverse selection problems virtually eliminate the possibilities to borrow against one's future labor income. This is also true for the framework presented here. Due to the separation of owner and service an agent who has sold his human capital would not receive a wage for contributing the service of human labor to the production process. Instead, the investor in the corresponding security earns a rental. Thus, after issuing the securities and receiving the proceeds there is a strong incentive to reduce the effort or not to work at all. More severe, investors might anticipate this and expect only persons to offer their human capital who have a higher tendency not to perform the contracted labor service as legal enforcement is virtually impossible. Hence, the market for human capital securities collapses and young agents are constraint from borrowing against their future wage income.

This imperfection has two major consequences for the young: First, although they still represent the future labor force and are in this sense still personally endowed with human capital, they cannot capitalize on this fact. Since a young agent has no sources of income other than borrowing against his human capital he has no actual consumption-savings decision. Without any wealth tradable against current consumption he is forced to save his entire endowment. Second, in his portfolio decision he cannot diversify his allocations into physical capital but is constraint to hold human capital only. This reduces the overlapping generations model to absurdity: the young have no role in the economy, as they can only act in their middle-age. More severe, without any

[1] Section 5.2.3 will elaborate further on this perspective.

possibility to consume one would expect the young to starve and the demographic structure to collapse. Doubtlessly, this represents major deterioration in the young's utility.[2]

The middle-aged are also heavily affected by the nontradability of human capital. Without prior investments in real capital they have only labor income and thus would be fully exposed to fluctuations in labor's share of output. On the other hand, in the savings for their old age they are limited to investments in physical capital securities. Being risk averse they suffer a loss in welfare as well. Finally, an old agent is negatively affected since he must draw on funds from savings allocated to physical capital only, i.e. a sub-optimally diversified portfolio.[3] To sum up, failure in the market for human capital securities worsens the situation of all agents.

Social Planer

To prevent this the social planer intervenes and adds further imperfections to the economy. In the spirit of Lipsey and Lancaster's [1956] second-best solutions, the idea is that the social planer can replicate the first-best solution established in Chap. 3 by correcting the imperfections through additional interferences. In this sense the planer's actions are Pareto-improving compared to the situation of failure in the market for human capital securities. The social planer has two means to influence the economy: he installs a pay-as-you-go pension system and introduces a general consumption tax. The PAYGO scheme is central for replicating tradable human capital. By paying old agents benefits funded by the middle aged, the old participate in labor's share of output – as they would had they invested in human capital securities. This is achieved because the contributions are proportional to the wages – as typically observed. In this sense the PAYGO system serves as redistribution mechanism transferring labor's share of output partially to the old generation. However, it also transfers the risk embedded in it!

As explained above, the young would starve when human capital is nontradable. Hence, the social planer also introduces a fiscal system which collects a general consumption tax and pays the revenues to the young. Out of these benefits they can consume and invest in physical capital. When middle-aged these investments give them a stake in physical capital's share of output – as they would have by diversification of their portfolio under tradable human capital.

Based on the World Bank's goals and the requirements set in Sect. 2.1, the

---

[2]Yet, one has to remember from Chap. 3 that three generations is the minimum number of cohorts that allows to capture the desired features for the pension problem. Simply, reducing the benchmark model to one of only two overlapping generations is thus not an option.

[3]Admittedly, there might be state realizations in which the old end up better off with participation in physical capital only. But since this is not a case of stochastic dominance it does not matter in an ex-ante consideration of expected utility.

PAYGO and the fiscal systems are supposed to be sustainable. The simplest implementation of this is to require balanced budgets in each scheme. Furthermore, the contribution as well as the tax rate are supposed to be constant over time. The social planer calibrates them in such a way that the resulting allocations perfectly match the first-best allocations of the unconstrained economy. Thus the central task of this chapter is to determine these rates. With them, the pay-as-you-go pension scheme and the consumption tax will allow to replicate exactly those allocations that would be realized in the perfect markets' case. Of course, in order to achieve robustness this replication must hold not only in the current period, but also in any future time step – regardless of the state of the economy then, i.e. regardless of the realizations of the stochastics!
The analysis will abstract from potential problems when the proposed systems are implemented. As such an implementation would likely happen gradually there could potentially be major distortions in this process. Instead, the analysis treats the replication as if it had always existed. Furthermore, it is assumed that all agents are fully aware of the established systems and account for them in their decisions. In other words, they do not consider a possible collapse of the public pension scheme and engage, for instance, in any extra precautionary savings for such a case.
The replication refers to the goods market only. Since only consumption enters the agents' utility they care only about this. The mere existence and tradability of human capital does not make them better off. Agents only benefit indirectly by the enhanced traded-off in the intertemporal consumption-savings decision enabled by the human capital securities. Therefore it is only necessary to achieve each agent's ability to consume the same amount as in the unconstrained case and that an identical amount of investments into physical capital is realized.[4] Portfolio allocations can and must obviously not be replicated in order to achieve the second-best solution. There would be no point in trying this without the human capital securities. Of course one could extend the definition of savings to include the implicit holding of PAYGO pension entitlements. However, for analyzing the implications – as will be done in Chap. 5 – it is easier to refer to the hypothetical first-best solution.

4.1.2 Outline of Replication

To avoid confusion of the replication with the complete markets' case a careful approach is chosen. Instead of only partially adjusting the model developed in Chap. 3, a second economy is established with the only purpose of mimicking the unconstrained results. This has the advantage that the consequences of

[4] Assuming that output is identical in the constraint and in the unconstrained case this last requirement is implied by market clearing. However, the condition that investments amount to the same size seems more logical, as only this can force output to mimic that of the complete markets' case.

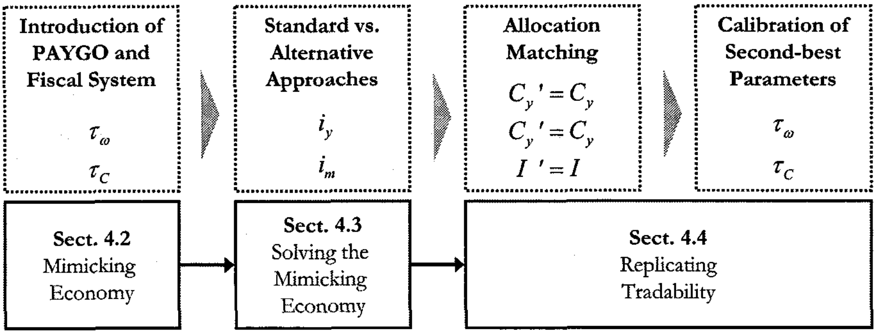

Fig. 4.1. Outline of replication

Remarks: Replication of human capital's tradability is shown in three steps. Firstly, a PAYGO and a fiscal system are added to the economic framework of the mimicking economy. Since the standard approaches of solving the intertemporal problem fail an alternative mechanism for solving the model is developed in the second step. Finally, this approach allows to calibrate the parameters of public interference so that the same allocations of consumption and investment are achieved as in the complete markets' case.

nontradable human capital as well as the conditions for the replication can be presented clearly and separately. For simplicity the notation in the mimicking economy follows the original one but with primes. Replication of the first-best results is then achieved in four major steps. These are indicated in Fig. 4.1. First, the framework of the mimicking economy is established by eliminating the human capital securities from the benchmark model. In order to thoroughly describe the consequences of this severe step, Sect. 4.2 reestablishes the model without the hypothetical securities but under the prime notation. To this setting, the pay-as-you-go and the fiscal system are added, characterized by their contribution rates τ_ω and τ_C. Both systems influence the budget constraints in the intertemporal utility maximization of all agents. The resulting substantial alterations present a major challenge for solving the model. These problems are indicated in Sect. 4.3 and cause the failure of the standard approaches, which allowed the straightforward procedure in the last chapter. Therefore an alternative characterization is developed in this step. This approach rests heavily on the underlying overlapping generations framework and its implications for aggregation. With this approach it is finally possible to use the conditions for a perfect matching of the complete markets' consumption and investment allocations. Section 4.4 does so and solves for the parameters of the PAYGO and fiscal system. In addition to this, it also relates the general results to the special case of Merton's [1983] analysis from which the calibration mechanism heavily draws. A short summary is given in Sect. 4.5.

4.2 Mimicking Economy

The setting for the replication with a PAYGO system and a consumption tax is essentially the same as described in Chap. 3 – apart from human capital being tradable. This section introduces the underlying macroeconomic framework and the social planer's interference by a PAYGO and a fiscal system. In order to distinguish the incomplete market variables from their perfect market correspondents a prime is added in the notation.[5] While one could principally distinguish the notation only for those variables where absolutely necessary a less parsimonious approach is chosen here. The mimicking economy including the public sector interventions is established independently first. Then the parameters for these interferences are determined so that the mimicking economy perfectly replicates the result of the Chap. 3. The separate prime notation allows to do this with intermingling the first- and the second-best model.

4.2.1 Maintained Setting

Major parts of the framework set up in Chap. 3 are identical for the mimicking economy. They are reintroduced here in order to clearly show where the changes are compared to the case of tradability.

Population

Like its complete markets' counterpart the mimicking economy is established in discrete time. In each period t, there are three living cohorts: young, middle-aged and old agents, with cohort sizes $N'_y(t)$, $N'_m(t)$, and $N'_p(t)$. The demographic development is again certain in the sense that the size of young generation is drawn from a random distribution, $\tilde{N}'_y(t) \in \mathbb{N}$, at the beginning of each period and that life-expectancy of three periods is certain:

$$N'_y(t) = \tilde{N}'_y(t)\,, \tag{4.2.1}$$

$$N'_m(t) = N'_y(t-1)\,, \tag{4.2.2}$$

$$\text{and} \quad N'_o(t) = N'_m(t-1) = N'_y(t-2)\,. \tag{4.2.3}$$

All agents within a cohort are absolutely homogeneous, without any idiosyncratic characteristics.

Consumption and Capital Technologies

The production of consumable output, $Y'(t)$, is again based on a Cobb-Douglas technology using physical capital, $K'(t)$, and human labor, $L'(t)$:

[5] Consequently, in this notation "$'$" is *not* used – as it frequently is – to denote a derivative.

$$Y'(t) = \tilde{A}'(t)\, K'(t)^{\tilde{\theta}'(t)}\, L'(t)^{1-\tilde{\theta}'(t)}, \tag{4.2.4}$$

where $\tilde{A}'(t)$ is the random total factor productivity and $\tilde{\theta}'(t)$ the uncertain production elasticity of physical capital with $0 \leq \tilde{\theta}'(t) \leq 1$. The realizations of both are revealed at the beginning of each period. Abstracting from monetary aspects in this real setup, the price of produced consumption goods is still normalized to one, $p'_Y(t) = 1$. These units can either be used for aggregate consumption, $C'(t)$, or investments in physical capital, $I'(t)$:

$$Y'(t) = C'(t) + I'(t). \tag{4.2.5}$$

In form of investments they allow the accumulation of physical capital based on a second technology of the Cobb-Douglas type as developed in great detail in Sect. 3.2.2. Hence, the physical capital accumulation is again characterized by

$$K'(t+1) = B'\, I'(t)^{\vartheta'}\, K'(t)^{1-\vartheta'}. \tag{4.2.6}$$

Consequently, the price of acquiring a unit of physical capital in terms of consumption goods, $p'_K(t)$, is given by the mimicking economy's equivalent to (3.2.11),

$$p'_K(t) = \frac{1}{B'\vartheta'} \left(\frac{I'(t)}{K'(t)} \right)^{1-\vartheta'}. \tag{4.2.7}$$

And the value of the entire real capital stock – of all physical capital units carried into the next period – is

$$K'(t+1) p'_K(t) = \frac{1}{\vartheta'} I'(t) \tag{4.2.8}$$

by analogy to (3.2.12). Similarly, each unit of physical capital earns again a rental of $(1-\vartheta')/\vartheta' I'(t)$ in the capital accumulation technology. As in (3.3.4) gross domestic product, $Y'^{GDP}(t)$, exceeds national income by this rental:

$$Y'^{GDP}(t) = C'(t) p'_Y(t) + K'(t+1) p'_K(t) = Y'(t) + \frac{1-\vartheta'}{\vartheta'} I'(t). \tag{4.2.9}$$

Firms and Factor Markets

Because the only imperfection concerns the mimicking economy's capital market, firms are again in perfect competition and maximize profits – without any agency problems. According to (4.2.6) decisions on the physical capital stock have to be made in period $t-1$ so that profit maximization is achieved by adjusting the demand for human labor:

$$\max_{L'(t)} \Pi'(t) = Y'(t) - \omega'(t)\, L'(t), \tag{4.2.10}$$

where $\omega'(t)$ is the wage rate. Since only the middle-aged generation works in the overlapping generations framework, equilibrium in the labor market is characterized by

$$L'(t) = N'_m(t) = N'_y(t-1). \tag{4.2.11}$$

Assuming competitive factor markets, labor's share of output is again

$$\omega'(t)L'(t) = \left(1 - \tilde{\theta}'(t)\right) Y'(t). \tag{4.2.12}$$

And the equilibrium wage rate is thus given by

$$\omega'(t) = \frac{(1 - \tilde{\theta}'(t))Y'(t)}{N'_m(t)}. \tag{4.2.13}$$

As in the standard functional distribution of income based on marginal productiveness, this implies that physical capital's share of output is $\tilde{\theta}'(t)Y'(t)$. Adding to this the rental it earns in the adjustment technology and dividing the sum by the size of the real capital stock gives the total rental per unit of physical capital:

$$\frac{\tilde{\theta}'(t)Y'(t)}{K'(t)} + \frac{1-\vartheta'}{\vartheta'}\frac{I'(t)}{K'(t)} = \frac{\vartheta'\tilde{\theta}'(t)Y'(t) + (1-\vartheta')I'(t)}{\vartheta' K'(t)}. \tag{4.2.14}$$

Financial Markets

Since in the mimicking economy the securities tracking the value of human capital do *not* exist, the capital market does only consist of the physical capital and the riskless one. The latter yields a predetermined return of $R'(t)$ and is in zero-net supply. Let each unit of the former represent ownership of one unit of the mimicking economy's physical capital, so that there is a total of $K'(t+1)$ physical capital securities. By analogy to (4.2.14) the return on them is

$$\tilde{R}'_K(t+1) = \frac{\tilde{\theta}'(t+1)\tilde{Y}'(t+1) + \frac{1-\vartheta'}{\vartheta'}\tilde{I}'(t+1)}{K'(t+1)p'_K(t)} \tag{4.2.15}$$

or using (4.2.8)

$$\tilde{R}'_K(t+1) = \frac{\vartheta'\tilde{\theta}'(t+1)\tilde{Y}'(t+1) + (1-\vartheta')\tilde{I}'(t+1)}{I'(t)}. \tag{4.2.16}$$

Without any further securities in positive supply (4.2.16) represents also the return on the capital market, i.e.

$$\tilde{R}'_M(t+1) = \tilde{R}'_K(t+1), \tag{4.2.17}$$

and (4.2.8) characterizes also the total market capitalization.

4.2.2 Public Sector

As briefly introduced in Sect. 4.1.1 there are two aspects of intervention by the social planer: a pay-as-you-go public pension scheme and a fiscal system. The former is the subject of interest here, an idealized implementation of a typical public pension scheme, the latter is a general consumption tax funding a subsidy to the young.

Pay-as-you-go Pension Scheme

In the mimicking economy there exists a pension system of the PAYGO-style, discussed in Sect. 2.2.6. The simple structure of the overlapping generations framework allows for a very idealized form of it: while young agents do not participate in this public pensions scheme, a middle-aged agent makes an obligatory contribution proportional to his wage. In exchange for this contribution he receives a benefit from PAYGO when he is old.

Because the purpose of this research are neither questions of implementing, i.e. phasing in, a public pension scheme, nor of bridging cohort fluctuations by running deficits, but the pure effect of PAYGO in a financial risk-return perspective, it is assumed that it maintains a *balanced* budget at all times. This assumption eliminates any form of "political" risk in its return: the governing body cannot deliberately set new rules to benefit particular interest groups by increasing their pension entitlements or protecting them from cuts. These aspects and their immediate consequence in reality – i.e. financing problems – are interesting problems per se but more likely to be addressed in an institutional analysis of actual PAYGO pension systems of different countries. In the approach here, which focuses on the economic efficiency argument, they would weaken the link of macroeconomic reality and implied returns in the pension system. In addition to this, many imbalances in the pension system stem from political aim of using the PAYGO system as a device for income redistribution. Again it is certainly an interesting question whether, for instance, educational years without actual contribution payments should count towards pension entitlement or how motherhood can be integrated into the system. However, the redistributive character of them distracts from the efficiency argument to be made here. Furthermore, the next paragraph will introduce a separate fiscal system, which is more suitable for dealing with redistributing interferences. Yet, note that implicit returns from PAYGO are not certain even if political influence on them is ruled out. On the contrary, they are exposed more to the underlying economic development, because there are no possibilities of intertemporal buffering by postponing disadvantageous consequences through deficits.

Turning to the formal analysis, let the size of the PAYGO's total budget in period t be $\mathfrak{T}_o(t)$.[6] It is financed by collecting contributions from the labor

[6] Since it is not present in the complete markets' case primes would be redundant for the public sector variables and are thus omitted. Furthermore note, that the index

force proportional to the wages. With a constant contribution rate of τ_ω this implies[7]

$$\mathfrak{T}_o(t) = \tau_\omega \times \omega'(t) \times L'(t) = \tau_\omega \times \omega'(t) \times N'_m(t), \tag{4.2.18}$$

because all members of the middle-aged generation are working as seen in (4.2.11). For the same reason and due to (4.2.3) *all* members of the old generation have paid into the system in the preceding period and are thus entitled with pension benefits. Using $\mathfrak{t}_o(t)$ to denote the benefits an individual retired agent receives, the assumption of a balanced budget implies

$$\mathfrak{T}_o(t) = \mathfrak{t}_o(t) \times N'_o(t). \tag{4.2.19}$$

Combining (4.2.18) and (4.2.19), one can establish

$$\begin{aligned} \mathfrak{t}_o(t) &= \tau_\omega\, \omega'(t) \frac{N'_m(t)}{N'_o(t)} \\ &= \tau_\omega\, \omega'(t) \frac{N'_y(t-1)}{N'_y(t-2)}. \end{aligned} \tag{4.2.20}$$

This result, familiar from public finance, relates benefits from a PAYGO system to the demographic development. Benefits to an old-aged agent depend positively on the contribution rate, the wage rate, the size of the working middle-aged cohort and negatively on the number of old-aged agents. Shifting the last result by one period and dividing by the contributions gives the implicit return on PAYGO, $\tilde{\mathfrak{R}}(t+1)$,

$$\begin{aligned} \tilde{\mathfrak{R}}(t+1) &= \frac{\tilde{\mathfrak{t}}_o(t+1)}{\tau_\omega \omega'(t)} = \frac{\tau_\omega\, \tilde{\omega}'(t+1) \frac{N'_y(t)}{N'_y(t-1)}}{\tau_\omega \omega'(t)} \\ &= \frac{\tilde{\omega}'(t+1)}{\omega'(t)} \frac{N'_y(t)}{N'_y(t-1)}. \end{aligned} \tag{4.2.21}$$

It depends on two components: the growth of the wage rate and the growth the young cohort's size. Due to the randomness of $\tilde{A}(t+1)$ and $\tilde{\theta}(t+1)$, period $t+1$'s output is uncertain and so is the wage rate as determined by (4.2.13). Therefore $\tilde{\mathfrak{R}}(t+1)$ is risky as well – even though deliberate political influence on it has been ruled out.

---

in the notation of the public sector budget follows the respective beneficiaries: "o" for the PAYGO system as the old receive the pensions and "y" for the tax system since only the young receive the proceeds.

[7]The assumption that the contribution rate is constant over time reflects the motive to rule out political influence on the system. Yet, there might be generalizations of the model in which the contribution rate is endogenously determined by other variables and is not constant but dynamically responding to changes in the economy and demography - but still not deliberately set by the governing body. In that sense a constant τ_ω is the simplest case of balanced PAYGO system.

Fiscal System

The social planer's second intervention in the competitive markets is by a fiscal system collecting a general tax on consumption. Its only purpose is to fund a transfer payment to the young generation. This subsidy is necessary as the young generation cannot capitalize on its endowment with human capital when human capital is not tradable. In the case of unconstrained, complete markets the young generation has been able to sell its human capital securities partially in order to diversify the savings and – more importantly – to consume out of the proceeds. As explained in Sect. 4.1.1, without the market for human capital securities the young cannot borrow against their future wages and would have to wait until being in the middle-age to live off their realized labor income. But if they cannot consume as young, one would expect that the model collapses since the young would not survive and the demographic structure implodes.

The reason for this problem is the simplicity of the model: it ignores the role of families as well as it embeds a very stylized life-cycle. Like in Constantinides et al. [2002] one could change the basic setup to include the young generation into the workforce and differentiate their labor efficiency from the middle-aged ones by some scaling factor. Then an agent would earn some wage already in his first period and transfer only the remaining capitalized value of human capital into the subsequent period, in which he works again. With his labor income partially earned in his youth, an agent could consume out of it and the problem of a collapsing demographic structure is mitigated. Or one could introduce families where middle-aged and old-aged agent transfer some consumption to their offspring, the young. Each of these ways to address the problem of a collapsing population would require *substantial* modifications of the framework.

The approach chosen here follows Merton [1983] and is a fiscal system providing benefits to the young by collecting a general consumption tax. With this exclusive expenditure any taxation of the young's consumption directly benefits themselves so that only the middle-aged and the old bear the tax. In this sense, it reflects somehow the family-idea: middle-aged and old devoting a part of their consumption to the raising of the young. Another justification is the observation that a substantial part of government expenditures is targeted towards children and the young generation:[8] kindergartens, public schooling and universities or direct subsidies for parenthood.[9] Thus the introduction of a fiscal system to benefit the young is not completely unrealistic even though its design will of course not cope with the complexity of real tax codes.

[8]However, in reality the increasing pension problem and a high debt burden have certainly diverted fiscal budgets from sufficient investments in and for future generations.

[9]Referring to the notion of redistributive goals within the PAYGO system another point here could be acknowledgment of motherhood years in the calculation of pension benefits, which is present in many systems.

The implementation is rather simple: there is a proportional consumption tax, τ_C, imposed on any unit of consumption, regardless of the consumer's age. The proceeds of this are paid to the period's young agents, each of whom receiving a benefit of $\mathfrak{t}_y(t)$. Requiring again a balanced budget at all times of size $\mathfrak{T}_y(t)$, payments to the young must equal the total of collected taxes:

$$\mathfrak{t}_y(t) \times N_y'(t) = \mathfrak{T}_y(t) = \tau_C \times C'(t)\,. \tag{4.2.22}$$

Thus, for consumption and diversification in savings each member of the young cohorts receives a transfer of

$$\mathfrak{t}_y(t) = \tau_C \frac{C'(t)}{N_y'(t)}. \tag{4.2.23}$$

4.2.3 Agents' Utility and Budget Constraints

The introduction of the pay-as-you-go scheme as well as of the fiscal system directly influence the agents' optimization. Furthermore, their utility maximization is seriously affected by the nontradability of human capital.

Utility

As established in Sect. 3.4.1 agents shall derive utility only out of consumption over the life-cycle – $c_y'(t)$, $c_m'(t+1)$ and $c_o'(t+1)$ – based on a time-separable utility function and a subjective discount factor β'. Limiting the analysis to the specific case of logarithmic utility and using the intuitive indexation in the three generations framework gives again

$$\begin{aligned} U_y'(t) &= \ln c_y'(t) + \mathbb{E}_t[\beta' \times \ln c_m'(t+1) + \beta'^2 \times \ln c_o'(t+2)] \\ &= \ln c_y'(t) + \mathbb{E}_t[\beta' \times U_m'(t+1)] \end{aligned} \tag{4.2.24}$$

for a young agent's utility,

$$\begin{aligned} U_m'(t) &= \ln c_m'(t) + \mathbb{E}_t[\beta' \times \ln c_o'(t+1)] \\ &= \ln c_m'(t) + \mathbb{E}_t[\beta' \times U_o'(t+1)] \end{aligned} \tag{4.2.25}$$

for a middle-aged agent, and

$$U_o'(t) = \ln c_o'(t) \tag{4.2.26}$$

for the utility of an old agent. Since agents derives utility exclusively from consumption there are again no bequests by the old.

Budget Constraints

Turning to the budget constraints, the convenient inclusion of labor income into the portfolio-payoff as a return on the human capital security is no longer possible. Instead, the standard approach of additive income sources must be applied. Furthermore, the PAYGO and fiscal system introduced in Sect. 4.2.2 must be addressed. Taken together these alterations imply that each agent has now principally three sources of funds: income from physical capital based on previous savings, labor income in the form of wages when the agent is working, and transfer income from the public pension scheme or from the fiscal system. The expenditure side must be supplemented by contributions to PAYGO and tax payments, too. Reflecting reality, however, contributions to the pension system will be subtracted directly from the wages and not accounted for as usage of funds. Furthermore, using a gross definition of consumption, i.e. including the tax attached to it, the definition of wealth from Chap. 3 is also applicable here: wealth $w'_a(t)$ of an agent aged a are all his sources of funds which he uses for gross consumption, $(1+\tau_C)c'_a(t)$ or savings $s'_a(t)$:[10]

$$w'_a(t) = (1+\tau_C)c'_a(t) + s'_a(t). \tag{4.2.27}$$

However, the intertemporal characterization of this wealth variable does not follow the standard stock-flow model of Chap. 3. Instead, the additional sources of funds and the PAYGO contributions must be addressed. Based on the explanations in Sect. 4.2.2, wealth of a young agent in the mimicking economy is given by the benefits he receives from the fiscal system

$$w'_y(t) = \mathfrak{t}_y(t). \tag{4.2.28}$$

A middle-aged agent receives income from previous savings, where he earns a gross return of $R'^{*}_{P,y}(t)$ on his portfolio,[11] and from working, where mandatory contributions to the PAYGO system are subtracted from his gross wage $\omega'(t)$. Thus he has in total

$$w'_m(t) = s'_y(t-1)R'^{*}_{P,y}(t) + (1-\tau_\omega)\omega'(t) \tag{4.2.29}$$

for gross consumption and saving. Finally, each old agent draws from his appreciated private savings as a middle-aged $s'_m(t-1)R'^{*}_{P,m}(t)$ and receives the public pension benefits from the PAYGO scheme:

$$w'_o(t) = s'_m(t-1)R'^{*}_{P,m}(t) + \mathfrak{t}_o(t). \tag{4.2.30}$$

Nota bene: equations (4.2.28) to (4.2.30) present a major extension to the definition of wealth as usually applied in finance. Wealth comprises *all* sources of net income and allows thus for the intuitive logic of saving equaling wealth minus gross consumption as shown in (4.2.27). In this sense, it represents a natural extension of the standard intertemporal approach in the spirit of Merton [1969*b*], Samuelson [1969], or Fama [1970*b*]. Apart from the interference

---

[10] Savings do thus only refer to the explicit forgoing of consumption to establish funded savings and do *not* include the implicit PAYGO savings.

[11] Admittedly the portfolio problem is trivial, but in order to facilitate comparison with Chap. 3 it is initially maintained.

of the public sector it addresses in particular the variable typically ignored in finance: labor income. From an economic point of view all forms of income are indistinguishable as they all allow for the only and ultimate goal of consumption, either immediately or postponed by savings. Therefore, the interpretation of wealth here does not only match the one from Chap. 3, but is also consistent with the explanations of (2.2.3) in Sect. 2.2.

4.2.4 Summary of Mimicking Economy

Section 4.2 has established the mimicking economy as base for the replication. This economy has emerged from the benchmark model by elimination of human capital's tradability. As surrogate the social planer has introduced a public sector comprising two compulsory systems. Table 4.1 briefly summarizes this setting in order to facilitate comparison with the complete markets' framework.
Even though the alterations compared to Chap. 3 look small – nontradability and a public sector – they have major consequences for the model and the solutions mechanism. Before digging into this, the next section will continue last chapter's numerical example of the model without real capital accumulation.

4.2.5 Full Depreciation and Numerical Example

In the benchmark model of tradable human capital the limiting case of fully depreciating physical capital turned out to be a simplifying version and was thus of special interest. Therefore it will again be analyzed separately and form again the theoretical foundation for the numerical illustration.

No Physical Capital Accumulation

The case without accumulation of physical capital is characterized again by $B' = 1$ and $\vartheta' = 1$. As in Chap. 3 the capital adjustment technology (4.2.6) reduces then to a trivial

$$K'(t+1)|_{\vartheta'=1} = I'(t)|_{\vartheta'=1}. \tag{4.2.31}$$

Existing physical capital fully depreciates and only new investments form the real capital stock of the subsequent period. The price of acquiring a unit of physical capital expressed in consumption goods is unity according to (4.2.7). Hence, the value of all physical capital units carried into the next period equals the value of investments

$$K'(t+1)|_{\vartheta'=1} \times p'_K(t)|_{\vartheta'=1} = I'(t)|_{\vartheta'=1}. \tag{4.2.32}$$

Concerning the public sector or the description of the agents' utility maximization there are no differences from the general case.

Table 4.1. Summary of mimicking economy

Implications for Budget Constraints

Young	$w'_y(t)$	$= \mathfrak{t}_y(t)$	(4.2.28)
Middle-aged	$w'_m(t)$	$= s'_y(t-1)R'^{*}_{P,y}(t) + (1-\tau_\omega)\omega'(t)$	(4.2.29)
Old	$w'_o(t)$	$= s'_m(t-1)R'^{*}_{P,m}(t) + \mathfrak{t}_o(t)$	(4.2.30)

Public Sector

PAYGO	Revenues	$\mathfrak{T}_o(t)$	$= \tau_\omega \times \omega'(t) \times N'_m(t)$	(4.2.18)
	Benefits	$\mathfrak{T}_o(t)$	$= \mathfrak{t}_o(t) \times N'_o(t)$	(4.2.19)
Fiscal System	Revenues	$\mathfrak{T}_y(t)$	$= \tau_C \times C'(t)$	(4.2.22)
	Benefits	$\mathfrak{T}_y(t)$	$= \mathfrak{t}_y(t) \times N'_y(t)$	(4.2.22)

Financial Markets

Physical Capital	Payoff		$\frac{\vartheta'\tilde{\theta}'(t)Y'(t)+(1-\vartheta')I'(t)}{\vartheta' K'(t)}$	(4.2.14)
	Return	$\tilde{R}'_K(t+1) =$	$\frac{\tilde{\theta}'(t+1)\tilde{Y}'(t+1)+\frac{1-\vartheta'}{\vartheta'}\tilde{I}'(t+1)}{K'(t+1)p'_K(t)}$	(4.2.15)
Human Capital	Payoff		$(1-\tilde{\theta}'(t))Y'(t)$	(4.2.12)
	Return	n.m.	Human capital is not tradable.	
Capital Market	Return	$\tilde{R}'_M(t+1) =$	$\frac{\tilde{\theta}'(t+1)\tilde{Y}'(t+1)+\frac{1-\vartheta'}{\vartheta'}\tilde{I}'(t+1)}{K'(t+1)p'_K(t)}$	(4.2.12)

Real Sector

Consumable Output	$Y'(t)$	$= \tilde{A}'(t) \times K'(t)^{\tilde{\theta}'(t)} \times L'(t)^{1-\tilde{\theta}'(t)}$	(4.2.4)
Capital Accumulation	$K'(t+1)$	$= B' \times I'(t)^{\vartheta'} \times K'(t)^{1-\vartheta'}$	(4.2.6)
Demographics	$N'_o(t+1)$	$= N'_m(t) = N'_y(t-1)$	(4.2.3)

Remarks: The table summarizes the setting of the mimicking economy in which human capital is not tradable but a pay-as-you-go public pension scheme and a consumption tax have been established. Reference to the respective equations are given.

Illustrative Example Economy

The numerical example of Chap. 3 will be extended in illustrate the explanations. In order to trim the numerics and anticipate the replicating assumption of Sect. 4.4.1 let the demographic structure of the mimicking economy be the same as in the original. Based on Sect. 3.6.5 the population consists thus of $N_y'(0) = 100$ young agents, $N_m'(0) = 125$ middle-aged agents and $N_o'(0) = 156$ old agents. Again the illustrative economy is characterized by full depreciation of physical capital. The price of consumption goods is again normalized, i.e. $p_Y'(t) = \$1\,\forall\, t$. Investments in real capital in the current period $t = 0$ shall be $I'(0) = \$200$ so that this constitutes the physical capital stock for $t = 1$. Of course, it must be shown later that this amount of investments is indeed realized given the social planer's interventions. The uncertainty in the macroeconomic production function's scale and elasticity parameter is characterized by the joint probability distribution of Table 4.2. Again it must be emphasized that identity to Table 3.3 is just to anticipate Sect. 4.4.1's assumption for replication and hence to simplify the example.

Table 4.2. Probability distribution in mimicking economy

	$\tilde{\theta}'(1) = 0.2$	$\tilde{\theta}'(1) = 0.4$	
$\tilde{A}'(1) = 10$	$P_1 = 21\%$	$P_3 = 39\%$	$P_{A'=10} = 60\%$
$\tilde{A}'(1) = 25$	$P_2 = 14\%$	$P_4 = 26\%$	$P_{A'=25} = 40\%$
	$P_{\theta'=0.2} = 35\%$	$P_{\theta'=0.4} = 65\%$	$\sum = 100\%$

Remarks: The table gives the joint probability distribution of the scale and elasticity parameter for the numerical example in the mimicking economy. For simplicity of the example it matches perfectly the numerics of Table 3.3.

Aggregate output in $t = 1$, $\tilde{Y}'(1)$, is then characterized by four possible states according to (4.2.4). Equation (4.2.16) translates them into four possible realizations for the return on physical capital:

$$\tilde{Y}'(1) = \begin{cases} \$1,148.70 & \text{with}\, P_1 \\ \$2,297.40 & \text{with}\, P_2 \\ \$1,319.51 & \text{with}\, P_3 \\ \$2,639.02 & \text{with}\, P_4 \end{cases} \quad \text{and} \quad \tilde{R}_K(1) = \begin{cases} 115\% & \text{with}\, P_1 \\ 230\% & \text{with}\, P_2 \\ 264\% & \text{with}\, P_3 \\ 528\% & \text{with}\, P_4 \end{cases}$$

In order to facilitate comparison these are the same as in Sect. 3.2.6. Furthermore, let the present realizations be $A'(0) = 20$ and $\theta'(0) = 0.4$. With an assumed physical capital stock of $K'(1) = \$251.89$ and the labor force of $L(0) = N_m'(0) = 125$ this gives an output in period 0 of

$$Y'(0) = \$\,20 \times 125^{0.4} \times 251.89^{1-0.4} \approx \$3,308.71,$$

which is the same as the finding in Sect. 3.6.5. With an elasticity parameter of $\theta'(0) = 0.4$ the shares of output paid to physical capital and human labor are

$$\begin{aligned} \theta'(0)Y'(0) &= \quad 0.4 \times \$3,308.71 = \$1,323.48 \\ \text{and} \quad (1-\theta'(0))Y'(0) &= (1-0.4) \times \$3,308.71 = \$1,985.23. \end{aligned}$$

According to (4.2.13) the real wage rate in period 0 is hence

$$\omega'(0) = \frac{\$1,985.23}{125} \approx \$15.88.$$

Public Sector

Since the public sector interferences will be constructed so that allocations identical to the complete markets case are realized, assume that aggregate consumption in period 0 is again $C'(0) = \$3.108, 71$. As will be verified later, let the replicating tax rate be $\tau_C = 9.0\%$. Based on (4.2.22) the total size of the fiscal budget is then[12]

$$\mathfrak{T}_y(0) = 9.0\% \times \$3,108.71 \approx \$279.35.$$

Excluding a fiscal deficit the transfer payment to each member of the young cohort is

$$\mathfrak{t}_y(0) = \frac{\$279.35}{100} \approx \$2.79.$$

Using (4.2.18) and an assumed contribution rate of $\tau_\omega|_{\vartheta'=1} = 73.2\%$[13] the initial budget of the pay-as-you-go system in the illustrative economy amounts to

$$\mathfrak{T}_o(0) = 73.2\% \times \$15.88 \times 125 \approx \$1,453.60.$$

The no-deficit assumption implies then that each retired agent receives benefits from the PAYGO pension scheme amounting to

$$\mathfrak{t}_o(0) = \frac{\$1,453.60}{156} \approx \$9.32.$$

Note that this setting means that the fiscal as well as the pay-as-you go pension system are absolutely balanced.

[12]Slight deviations in the digits are again due to preceeding rounding errors.

[13]Indeed this figure is relatively high compared to observed contributions rates! Yet, as will be seen later, the specific approach chosen here implies exactly this number. For a detailed discussion and critical appreciation of the results and their implications see Chap. 5.

Budget Constraints

Above results and the definitions of Sect. 4.2.3 imply that a young agent in the illustrative example economy has a wealth of $w'_y(0) = \$2.79$ for consumption and saving. Let income from previous savings, which each middle-aged agent receives, be $s'^*_y(-1)R'^*_{P,y}(-1) = \$2.84$. Adding the net wage he earns in the production of output gives based on (4.2.29)

$$w'_m(0) = \$2.84 + (1 - 73.2\%) \times \$15.88 \approx \$7.09$$

as wealth budget. Thus, let an old agent's income from funded savings be $s'^*_m(-1)R'^*_{P,m}(-1) = \$6.21$. Together with his transfer income from PAYGO this gives him a budget of

$$w'_o(0) = \$6.21 + \$9.32 \approx \$15.53.$$

It is crucial to note that these assumptions on income from funded savings perfectly add up to the total income paid to physical capital; i.e. the two generations, the middle-aged getting $s'^*_y(-1)R'^*_{P,y}(-1)$ and the old getting $s'^*_m(-1)R'^*_{P,m}(-1)$, share together exactly the total payoff of real capital:

$$\$2.84 \times 125 + \$6.21 \times 156 \approx \$1,323.48.$$

This is the same figure as just established.

4.3 Solving the Mimicking Economy

So far the underlying framework of the mimicking economy has been established by withdrawing the hypothetical human capital securities and adding the pay-as-you-go scheme and tax system to the benchmark model. Now one would ideally like to proceed in the same way as in the complete markets' and simply solve the model straightforwardly. However, the sequential approach applied in Chap. 3 – deriving the optimality conditions by backward induction first, aggregating them to establish equilibrium in capital market then, and finally deducing general equilibrium – is *not* possible here. The additional elements in the agents budget constraints let the problem of non-additivity reemerge. To illustrate this major obstacle the failure of the standard techniques for solving the intertemporal problem is shown in this chapter first. The remainder of it is then devoted to the development of a solution mechanism which will later allow to solve for the parameters of the social planer's interventions. In order to do so, an alternative characterization of the economy is established that rests heavily on the simplicity of the demographic structure. Adding the requirements of balanced budgets allows finally to derive an integrated condition, which can be interpreted in a non-standard application of the stochastic discount factor approach. In Sect. 4.4 this alternative representation will be the starting point for calibrating the public sector systems to perfectly replicate the unconstrained economy's allocations.

4.3.1 Failure of Standard Approaches

Even though they fail, the standard approaches of solving the model are presented first in order to foster the understanding for the need of an alternative solution mechanism and fully appreciate it later. Furthermore, since the consumption-savings problems are very similar to those in the unconstrained case, it is natural to try the standard approach first. To illustrate the problems encountered three variations are shown: firstly, the normal backward induction by straightforward dynamic programming, then the same with a redefinition of the savings portfolio, and finally within the stochastic discount factor approach. Neither of them is suitable for solving the model.

Backward Induction

As the agents' policies are again such that decisions of period t depend only on information available in that period, backward induction by applying the method of stochastic programming is the first candidate to solve the problem. Let thus indirect utility of an a-aged agent be given by

$$J'_a[w'_a(t), t] \equiv \max\{U'_a(t)\}, \tag{4.3.1}$$

where $w'_a(t)$ is the agent's wealth when in generation a and $U'_a(t)$ his lifetime utility as defined in (4.2.24) to (4.2.26). Without bequests the old consume their entire wealth, so that this first step in the backward inductions yields

$$c'^*_o(t) = \frac{1}{1+\tau_C} w'_o(t). \tag{4.3.2}$$

Like any other agent an old one must pay the consumption tax in the mimicking economy. Hence, his net consumption is smaller than his wealth by a factor of $1/(1+\tau_C) < 1$. Eliminating the consumption tax by setting $\tau_C = 0$ would give (3.4.12) from the case with tradable human capital. Combing (4.2.27) with (4.3.2) implies that an old agent does not save:

$$s'^*_o(t) = 0. \tag{4.3.3}$$

Based on (4.3.1) indirect utility of an old agent is thus

$$J'_o[w'_o(t), t] = \ln\left(\frac{1}{1+\tau_C} w'_o(t)\right) = -\ln(1+\tau_C) + \ln w'_o(t). \tag{4.3.4}$$

Having solved the first step of the backward induction, one continues with searching for the optimal policy for the middle-aged. Indirect utility of such an agent is given by (4.2.25) and (4.3.1) as

$$J'_m[w'_m(t), t] = \max_{\substack{c'_m(t),\\ \alpha'_{K,m}(t),\\ \alpha'_{H,m}(t)}} \left\{\ln c'_m(t) + \mathbb{E}_t\left[\beta' U'_o(t+1)\right]\right\}. \tag{4.3.5}$$

Applying the technique of stochastic dynamic programming and inserting (4.3.4) this can be restated as

$$J'_m[w'_m(t),t] = \max_{\bullet}\left\{\ln c'_m(t) - \beta' \ln(1+\tau_C) + \beta'\,\mathbb{E}_t\left[\ln \tilde{w}'_o(t+1)\right]\right\}. \quad (4.3.6)$$

Using then the conditions (4.2.27) and (4.2.30) to replace $\tilde{w}'_o(t+1)$, one obtains

$$J'_m[w'_m(t),t] = \max_{\bullet}\left\{\begin{array}{l} \ln c'_m(t) - \beta' \ln(1+\tau_C) \\ + \beta'\mathbb{E}_t\left[\ln\left(\begin{array}{l}[w'_m(t) - (1+\tau_C)c'_m(t)] \\ \times \tilde{R}'_{P,m}(t+1) + \tilde{\mathfrak{t}}_o(t+1)\end{array}\right)\right]\end{array}\right\}. \quad (4.3.7)$$

Comparing this with (3.4.16) from the benchmark case, the effect of the PAYGO benefits the old receive is evident: the intertemporal budget constraint has an extra stochastic term $\tilde{\mathfrak{t}}_o(t+1)$, as future benefits depend on the future wage level according to (4.2.20). Trying to derive the optimal consumption policy by establishing the first order condition of the maximization problem yields

$$\frac{1}{c'_m(t)} + \beta'\mathbb{E}_t\left[\frac{-(1+\tau_C)\tilde{R}'_{P,m}(t+1)}{[w'_m(t) - (1+\tau_C)c'^*_m(t)]\,\tilde{R}'_{P,m}(t+1) + \tilde{\mathfrak{t}}_o(t+1)}\right] = 0, \quad (4.3.8)$$

which is *not* solvable for the control $c'^*_m(t)$ without additional assumptions because of the additive term in the denominator! The uncertain future PAYGO benefits, $\tilde{\mathfrak{t}}_o(t+1)$, are not only the single most important feature of the mimicking economy but also the major obstacle for deriving a solution of it. In other words, this first variation of the standard solution approach fails.

Intertemporal Transfer Return

Hence, one could try to eliminate the additional term in the budget constraint and include the PAYGO benefits artificially in the portfolio. Such a portfolio would capture the entire payoff of funded – by capital market investments – and unfunded – by PAYGO participation – transfers of present unconsumed wealth into the next period. Define therefore the return on this *artificial* pension-portfolio as

$$\tilde{R}'_{T,m}(t+1) \equiv \frac{[w'_m(t) - (1+\tau_C)c'_m(t)]\,\tilde{R}'_{P,m}(t+1) + \tilde{\mathfrak{t}}_o(t+1)}{w'_m(t) - (1+\tau_C)c'_m(t)}. \quad (4.3.9)$$

This definition allows to reformulate (4.3.7) as

$$\begin{aligned} J'_m[w'_m(t),t] &= \max_{\bullet}\left\{\begin{array}{l} \ln c'_m(t) - \beta' \ln(1+\tau_C) \\ + \beta'\mathbb{E}_t\left[\ln\left([w'_m(t) - (1+\tau_C)c'_m(t)]\,\tilde{R}'_{T,m}(t+1)\right)\right]\end{array}\right\} \\ &= \max_{\bullet}\left\{\begin{array}{l} \ln c'_m(t) + \beta' \ln[w'_m(t) - (1+\tau_C)c'_m(t)] \\ - \beta' \ln(1+\tau_C) + \beta'\mathbb{E}_t\left[\ln \tilde{R}'_{T,m}(t+1)\right]\end{array}\right\}, \quad (4.3.10)\end{aligned}$$

which seems to be a standard intertemporal problem. However, because the definition in (4.3.9) implies

$$\frac{\partial \tilde{R}'_{T,m}(t+1)}{\partial c'_m(t)} = \frac{(1+\tau_C)\tilde{\mathfrak{t}}_o(t+1)}{[w'_m(t) - (1+\tau_C)c'_m(t)]^2},$$

the problem is only shifted to the first order condition of the maximization in (4.3.10):

$$\frac{1}{c'^*_m(t)} - \frac{\beta'(1+\tau_C)}{w'_m(t) - (1+\tau_C)c'^*_m(t)}$$
$$+\beta'\mathbb{E}_t\left[\frac{1}{\tilde{R}'^*_{T,m}(t+1)}\frac{(1+\tau_C)\tilde{\mathfrak{t}}_o(t+1)}{[w'_m(t) - (1+\tau_C)c'_m(t)]^2}\right] = 0.$$

Again the non-linearity impedes further simplifications. Thus the idea of an aggregate intertemporal transfer return does not help with the problem either. Trying to address the non-linearity by redefining the portfolio definition to include the benefits from the pay-as-you-go scheme fails as well.

Stochastic Discount Factor

Possibly, asset pricing's standard approach of stochastic discount factors can overcome the hurdle. While (4.3.9) might have seemed to be an ad-hoc definition, it is directly linked to the SDF framework. For a middle-aged (4.2.25) implies that the SDF is given by

$$\tilde{M}'_m(t+1) = \beta'\frac{c'^*_m(t)}{\tilde{c}'^*_o(t+1)}. \tag{4.3.11}$$

Applying (4.3.2) and (4.2.29) this can be written explicitly as

$$\begin{aligned}\tilde{M}'_m(t+1) &= \beta'\frac{c'^*_m(t)}{\frac{1}{1+\tau_C}w'^*_o(t+1)} \\ &= \beta'\frac{(1+\tau_C)c'^*_m(t)}{[w'_m(t) - (1+\tau_C)c'^*_m(t)]\,\tilde{R}'^*_{P,m}(t+1) + \tilde{\mathfrak{t}}_o(t+1)}.\end{aligned} \tag{4.3.12}$$

Using now the intertemporal transfer return as defined in (4.3.9) allows to restate the SDF of a middle-aged agent as

$$\begin{aligned}\tilde{M}'_m(t+1) &= \beta'\frac{(1+\tau_C)c'^*_m(t)}{[w'_m(t) - (1+\tau_C)c'^*_m(t)]\,\tilde{R}'_{T,m}(t+1)} \\ &= \beta'\frac{(1+\tau_C)c'^*_m(t)}{s'^*_m(t)\,\tilde{R}'_{T,m}(t+1)}.\end{aligned} \tag{4.3.13}$$

Even though tax-distorted, the similarity of the SDF in the mimicking economy and the benchmark model is striking. $\tilde{R}'_{T,m}(t+1)$ has almost the role of the optimal portfolio return in the case with tradable human capital. This underpins its interpretation as aggregate intertemporal transfer return reflecting the portfolio of unfunded and funded pension vehicles. However, as the attempt to derive an optimal consumption policy with the intertemporal transfer return has not solved the problem of non-linearity in the budget constraint, the SDF approach – as established as it is – cannot take one any further either.

Furthermore, the last step in the backwards induction to solve the dynamic programming problem has not even been addressed yet. The intertemporal budget constraint from young to middle-aged includes another non-linearity. Combining (4.2.27) to (4.2.29) implies

$$\tilde{w}_m'(t+1) = \left(\mathrm{t}_y(t) - c_y'(t)\right) \tilde{R}^*_{P',y}(t+1) + (1-\tau_\omega)\tilde{\omega}'(t+1). \quad (4.3.14)$$

Again an additive term, $(1-\tau_\omega)\tilde{\omega}'(t+1)$, distorts the usual relationship. Yet, this uncertain future wage net of contributions to the PAYGO system reflects the intended framework twofold: Firstly, it address the standard design of public pension schemes by directly charging contributions from wages. And secondly, the macroeconomic framework introduces uncertainty about the wage level per se. Neither of these aspects can be neglected without giving up the requirement of consistency set in Sect. 2.1.

4.3.2 Aggregation and Capital Market Participation

Because the usual solution mechanism has failed – even in variations of it – an alternative approach to establish equilibrium in the mimicking economy is required. Following Merton [1983], the key trick is exploiting the simplicity of the population structure. This allows to derive important relations on an aggregate cohort-level, even though the optimal policies have not been solved for individual agents. They enter only implicitly.

Aggregating Assumptions

In order to use the demographic structure, cohort-aggregate variables must be defined first. This can be done analogously to the unconstrained case. Hence, consumption, saving and wealth of the three generations is given by

$$\begin{aligned} W_y'(t) &\equiv N_y'(t)w_y'(t), & C_y'(t) &\equiv N_y'(t)c_y'^*(t), & S_y'(t) &\equiv N_y'(t)s_y'^*(t), \\ W_m'(t) &\equiv N_m'(t)w_m'(t), & C_m'(t) &\equiv N_m'(t)c_m'^*(t), & S_m'(t) &\equiv N_m'(t)s_m'^*(t), \\ W_o'(t) &\equiv N_o'(t)w_o'(t), & C_o'(t) &\equiv N_o'(t)c_o'^*(t) & & \end{aligned} \quad (4.3.15)$$

with the cohort sizes, $N_y'(t)$, $N_m'(t)$ and $N_o'(t)$ as established in Sect. 4.2.1. The mimicking economy's aggregate variables are accordingly given by

$$\begin{aligned} W'(t) &\equiv W_y'(t) + W_m'(t) + W_o'(t), \\ C'(t) &\equiv C_y'(t) + C_m'(t) + C_o'(t), \\ \text{and} \quad S'(t) &\equiv S_y'(t) + S_m'(t). \end{aligned} \quad (4.3.16)$$

Participation in Capital Market

Inspecting now the capital market's setting outlined in Sect. 4.2.1 allows to make a first simplification: with only two kinds of securities the agents' portfolio problems is trivial. Since the securities tracking the capitalized value of human capital do not exist in the mimicking economy, all investors' allocation to it must be zero, i.e. $\alpha'_{H,a}(t) = 0$. As the riskless asset is again in zero net supply this translates into $\alpha'_{K,a}(t) = 1 - \alpha'_{H,a}(t) = 1$. Thus the portfolio return of an a-aged agent, $R'^{*}_{P,a}(t)$, defined in analogy to (3.4.21), can be reduced to[14]

$$\begin{aligned}\tilde{R}'^{*}_{P,a}(t+1) \equiv \;& \alpha'^{*}_{K,a}(t)[\tilde{R}'_K(t+1) - R'(t)] \\ & + \alpha'^{*}_{H,a}(t)[\tilde{R}'_H(t+1) - R'(t)] + R'(t) \\ = \;& \tilde{R}'_K(t+1) \qquad \text{for } a = y, m. \end{aligned} \tag{4.3.17}$$

The return on the portfolio of funded savings equals the return on physical capital. Furthermore observe that absence of the human capital security and the zero net-supply of the riskless asset also implies that the market portfolio consists only of physical capital:

$$\alpha'^{M}_{K}(t) = 1 \quad \text{and} \quad \alpha'^{M}_{H}(t) = 0. \tag{4.3.18}$$

With savings only realizable as investments in physical capital, a further distinction of them is not necessary. Yet, for the purpose of later analogy define savings allocated to physical capital, $S'_K(t)$, by

$$S'_K(t) \equiv S'(t) \times \alpha'^{M}_{K}(t). \tag{4.3.19}$$

Using (4.3.16) and (4.3.18) in (4.3.19) one can establish

$$S'_K(t) = S'_y(t) + S'_m(t). \tag{4.3.20}$$

Savings in physical capital are based on either savings of the young or savings of the middle-aged. Similar to the benchmark model's case, these savings dedicated to physical capital must match the value of physical capital carried into the next period, i.e.

$$S'_K(t) = K'(t+1)p'_K(t). \tag{4.3.21}$$

The interpretation of (4.3.21) becomes again clearer by addressing the firms' financing decision: in order to finance next period's physical capital stock with a value of $K'(t+1)p'_K(t)$ the firms issue $K'(t+1)$ securities at a price of $p'_K(t)$.

[14]Nota bene: while this simplification could have been introduced in Sect. 4.3.1, it would not have helped with the problem of non-linearity in the budget constraint. The triviality of the portfolio problem would have been reflected in the first order condition with respect to portfolio allocations, which have not been discussed in order to not distract from the non-linearity problem.

This supply must match investors' demand for physical capital securities, $S'_K(t)$. Again the adjustment cost assumption allows to link the market value of the physical capital to the amount of current investments. Using (4.2.8) one can reformulate (4.3.21) to obtain

$$S'_K(t) = S'_y(t) + S'_m(t) = \frac{1}{\vartheta'} I'(t) . \tag{4.3.22}$$

Now define the fraction of physical capital investment realized by the young and the middle-aged generations as

$$i'_y(t) \equiv \frac{\vartheta' S'_y(t)}{I'(t)} \tag{4.3.23}$$

$$\text{and} \quad i'_m(t) \equiv \frac{\vartheta' S'_m(t)}{I'(t)} , \tag{4.3.24}$$

where the savings in the nominator must be corrected by the parameter ϑ' to address the adjustment costs. With these definitions (4.3.22) can be written as

$$i'_y(t) + i'_m(t) = 1 . \tag{4.3.25}$$

Investments in physical capital are either carried out by the young or by the middle-aged generation. Usage of these fractions of investments and thus of physical capital ownership is the central trick to circumvent the non-linearity problem encountered in Sect. 4.3.1. The stylized design of the population structure with only three overlapping generations allows to conclude that all physical capital stock must be owned by either of the generations who save. Based on this insight an alternative formulation of the intertemporal budget constraint can be established.

4.3.3 Alternative Characterization

Exploiting these definitions allows to substantially reformulate the maximization problem. Using (4.3.23) and (4.3.24), the intertemporal constraints established in Sect. 4.2.3 can be rewritten for all agents.

Young Agents

Starting with the young observe that (4.3.23) implies $S'_y(t) = 1/\vartheta' \, i'_y(t) I'(t)$. By undoing the aggregation of (4.3.15) one obtains

$$s'^*_y(t) = \frac{S'_y(t)}{N'_y(t)} = \frac{1}{\vartheta'} i'_y(t) \frac{I'(t)}{N'_y(t)} . \tag{4.3.26}$$

Then the budget constraint and the specification of a young's wealth allow to characterize a young agent's optimal consumption contingent on $i'_y(t)$ as well. From (4.2.27) and (4.2.28) optimal consumption is given by

$$c'^*_y(t) = w'_y(t) - s'^*_y(t) = \mathfrak{t}_y(t) - \frac{1}{\vartheta'} i'_y(t) \frac{I'(t)}{N'_y(t)} . \tag{4.3.27}$$

Middle-aged Agents

In his middle age an agent participates in the returns of the capital market according to his investments as a young. The absolute size of this investment is given by (4.3.26) for $t-1$. The current value of the portfolio equals the product of the exposure and the return on the optimal portfolio, $R'^{*}_{P,y}(t)$. According to (4.3.17) the portfolio return is trivial being the same as the return on physical capital, $R'_K(t)$. Hence, it can be replaced by (4.2.16). Altogether one has

$$\begin{aligned} s'^{*}_y(t-1)R'^{*}_{P,y}(t) &= \frac{1}{\vartheta'}i'_y(t-1)\frac{I'(t-1)}{N'_y(t-1)}R'_K(t) \\ &= \frac{1}{\vartheta'}i'_y(t-1)\frac{I'(t-1)}{N'_y(t-1)}\frac{\vartheta'\theta'(t)Y'(t)+(1-\vartheta')I'(t)}{I'(t-1)} \\ &= i'_y(t-1)\left[\theta'(t)Y'(t)+\frac{1-\vartheta'}{\vartheta'}I'(t)\right]\frac{1}{N'_y(t-1)}. \end{aligned} \tag{4.3.28}$$

This is the income from real capital received by an individual middle-aged agent due to his savings when he was young: the first two factors reflect the middle-aged generation's stake in real capital's payoff; the fraction at the end corrects this by the number of agent in the cohort. Using this and the wage rate from (4.2.13) in the corresponding budget constraint, (4.2.29), allows to restate it as

$$\begin{aligned} w'_m(t) = i'_y(t-1)\left[\theta'(t)Y'(t)+\frac{1-\vartheta'}{\vartheta'}I'(t)\right]\frac{1}{N'_y(t-1)} \\ +(1-\tau_\omega)\frac{(1-\tilde{\theta}'(t))Y'(t)}{N'_m(t)}. \end{aligned} \tag{4.3.29}$$

Exploiting the demographic structure (4.2.2) and (4.2.27) for the usage of this wealth yields that savings of a middle-aged agent must equal

$$\begin{aligned} s'_m(t) = \frac{i'_y(t-1)}{N'_m(t)}\left[\theta'(t)Y'(t)+\frac{1-\vartheta'}{\vartheta'}I'(t)\right] \\ +(1-\tau_\omega)(1-\theta'(t))\frac{Y'(t)}{N'_m(t)} - (1+\tau_C)c'_m(t). \end{aligned} \tag{4.3.30}$$

Of course this must hold at the optimal policy, i.e. with $s'^{*}_m(t)$ and $c'^{*}_m(t)$, as well.

Old Agents

Like for the young, a middle-aged agent's savings give him a participation in the return of the capital market later. Hence, the fraction of income from physical capital received by an old agent due to his savings when middle-aged can be established in analogy to (4.3.28):

$$s'^{*}_{m}(t-1)R'^{*}_{P,m}(t) = i'_m(t-1)\left[\theta'(t)Y'(t) + \frac{1-\vartheta'}{\vartheta'}I'(t)\right]\frac{1}{N'_m(t-1)} \quad (4.3.31)$$

Using this with the population dynamics of (4.2.3), $N'_m(t-1) = N'_o(t)$, in the budget constraint of an old, (4.2.30), reads

$$w'_o(t) = \frac{i'_m(t-1)}{N'_o(t)}\left[\theta'(t)Y'(t) + \frac{1-\vartheta'}{\vartheta'}I'(t)\right] + \mathfrak{t}_o(t). \quad (4.3.32)$$

Because an old agent does not save, his optimal consumption is given by[15]

$$c'^{*}_o(t) = \frac{1}{1+\tau_C}\left[\frac{i'_m(t-1)}{N'_o(t)}\left(\theta'(t)Y'(t) + \frac{1-\vartheta'}{\vartheta'}I'(t)\right) + \mathfrak{t}_o(t)\right]. \quad (4.3.33)$$

He consumes his entire payoff from the capital market as well as his PAYGO benefits but must pay the consumption tax for each unit consumed.

4.3.4 Sustainability and Integration

Having derived alternative characterizations for each generation in the last section, these can be integrated further to a single equation. For this step the no-deficit assumption for sustainability of the pay-as-you-go scheme is incorporated.

Sustainable PAYGO

Using the wage rate of (4.2.13) in the condition for a balanced public pension scheme, (4.2.20), implies that PAYGO benefits are given by

$$\begin{aligned}\mathfrak{t}_o(t) &= \tau_\omega\, \omega'(t)\frac{N'_m(t)}{N'_o(t)} = \tau_\omega\, \frac{(1-\tilde{\theta}'(t))Y'(t)}{N'_m(t)}\frac{N'_m(t)}{N'_o(t)} \\ &= \tau_\omega\, \frac{(1-\tilde{\theta}'(t))Y'(t)}{N'_o(t)}. \end{aligned} \quad (4.3.34)$$

Equation (4.3.34) is a simple result with important implications: with a fixed contribution rate, τ_ω, pensions from a pay-as-you-go system, $\mathfrak{t}_o(t)$, depend

[15]Remember that this initial step of the traditional backward induction had not been affected by the non-linearity in the constraint. Hence, Sect. 4.3.1 yielded a result at least for this stage of the problem: the old dedicate their entire wealth to gross consumption.

positively on labor's share of output, $(1 - \tilde{\theta}'(t))Y'(t)$, and negatively on the number of retirees, $N_o'(t)$. In other words, when total output decreases – say due to a decreasing population – and an increase in the fraction of output paid to human labor does not compensate this, pension benefits per retiree have got to decline in order to maintain a balanced budget. The reduction is even more pronounced when the number of PAYGO recipients increases.

Alternative Intertemporal Budget Constraint

Plugging this condition on a balanced PAYGO budget and the notion from (4.3.24) and (4.3.15) that

$$\frac{i_m'(t-1)}{N_o'(t)} = \frac{\frac{\vartheta' S_m'(t-1)}{I'(t-1)}}{N_m'(t-1)} = \frac{\frac{\vartheta' S_m'(t-1)}{N_m'(t-1)}}{I'(t-1)} = \frac{\vartheta' s_m'^*(t-1)}{I'(t-1)}$$

into the alternative consumption constraint of an old agent, (4.3.33), reads

$$c_o'^*(t) = \frac{1}{1+\tau_C}\left[\begin{array}{l}\frac{\vartheta' s_m'^*(t-1)}{I'(t-1)}\left(\theta'(t)Y'(t) + \frac{1-\vartheta'}{\vartheta'}I'(t)\right)\\ + \frac{\tau_\omega(1-\theta'(t))Y'(t)}{N_o'(t)}\end{array}\right]. \tag{4.3.35}$$

This equation is a first step towards solving the mimicking economy: it allows to link an old agent's consumption to the savings of a middle-aged, where the fractional ownership of physical capital by the young has been integrated! To see this observe that (4.3.35) depends on $s_m'^*(t-1)$; yet, $s_m'^*(t)$ as formulated in (4.3.30) is a function of $i_y'(t-1)$.
Hence, shifting the condition on $c_o'^*(t)$ by one period and inserting (4.3.30) for $s_m'^*(t)$ gives a new intertemporal constraint for a middle-aged:

$$\tilde{c}_o'^*(t+1) = \frac{1}{1+\tau_C}\left\{\begin{array}{l}\left[\begin{array}{l}\frac{i_y'(t-1)}{N_m'(t)}\left(\theta'(t)Y'(t) + \frac{1-\vartheta'}{\vartheta'}I'(t)\right)\\ + (1-\tau_\omega)(1-\theta'(t))\frac{Y'(t)}{N_m'(t)} - (1+\tau_C)c_m'(t)\end{array}\right]\\ \times \frac{\vartheta'}{I'(t)}\left(\tilde{\theta}'(t+1)\tilde{Y}'(t+1) + \frac{1-\vartheta'}{\vartheta'}I'(t+1)\right)\\ + \frac{\tau_\omega(1-\tilde{\theta}'(t+1))\tilde{Y}'(t+1)}{N_o'(t+1)}\end{array}\right\}. \tag{4.3.36}$$

This alternative characterization is the key mechanism for solving the mimicking economy. The first line is a middle-agent's income from physical capital; the second adds his net labor income and subtracts his gross consumption expenditure so that the the first two lines represents the funded savings of that agent. Based on real capital's rental, these savings will earn a gross return, which is reflected by the third line's multiplication factor. For the resulting old-agent's wealth, one must finally add the future PAYGO benefits, which are characterized by the last item of (4.3.36).

SDF Interpretation

Using this alternatively formulated constraint one can revisit the intertemporal optimization problem of a middle-aged agent. According to (4.2.25) he maximizes

$$U'_m(t) = \ln c'_m(t) + \beta' \mathbb{E}_t \left[\ln \tilde{c}'_o(t+1)\right]$$

subject to the alternative budget constraint of (4.3.36).
Inserting the constraint into the objective function and taking the derivative with respect to the middle-aged's control, $c'_m(t)$, gives the first order condition for this problem:

$$\begin{aligned}\frac{\partial U'_m(t)}{\partial c'_m(t)} &= \frac{1}{c'_m(t)} + \beta' \mathbb{E}_t \left[\frac{1}{\tilde{c}'_o(t+1)} \frac{\partial \tilde{c}'_o(t+1)}{\partial c'_m(t)}\right] \\ &= \frac{1}{c'_m(t)} + \beta' \mathbb{E}_t \left[\frac{-1}{\tilde{c}'_o(t+1)} \frac{\vartheta' \left(\tilde{\theta}'(t+1)\tilde{Y}'(t+1) + \frac{1-\vartheta'}{\vartheta'} I'(t+1)\right)}{I'(t)}\right] \\ &= 0. \end{aligned} \tag{4.3.37}$$

Reorganizing the last line, shows that at the optimum it holds that

$$1 = \mathbb{E}_t \left[\beta' \frac{c'^*_m(t)}{\tilde{c}'^*_o(t+1)} \frac{\vartheta' \tilde{\theta}'(t+1)\tilde{Y}'(t+1) + (1-\vartheta')\tilde{I}'(t+1)}{I'(t)}\right]. \tag{4.3.38}$$

Using (4.3.11) and noting that the second fraction can be replaced by $\tilde{R}'_K(t+1)$ according to (4.2.16) allows the SDF-formulation to reemerge from (4.3.38):

$$1 = \mathbb{E}_t \left[\tilde{M}'_m(t+1)\tilde{R}'_K(t+1)\right]. \tag{4.3.39}$$

Consequently, the alternative formulation of the intertemporal constraint is consistent with the standard optimization rationale. Equation (4.3.36) reflects the intertemporal trade-offs correctly, yet in a non-standard way.

4.3.5 Full Depreciation and Numerical Example

Since the illustrative economy is based on the case of fully depreciating physical capital the specification of the alternative intertemporal budget constraint must be derived first. By plugging $\vartheta' = 1$ into (4.3.36) the alternative budget constraint of a middle-aged agent reduces for this special case to[16]

$$\tilde{c}_o'^*(t+1) = \frac{1}{1+\tau_C|_{\vartheta'=1}} \left\{ \begin{array}{l} \left[\begin{array}{l} \frac{i_y'(t-1)|_{\vartheta'=1}}{N_m'(t)} \theta'(t) Y'(t) \\ + (1-\tau_\omega|_{\vartheta'=1})(1-\theta'(t)) \frac{Y'(t)}{N_m'(t)} \\ - (1+\tau_C|_{\vartheta'=1}) c_m'(t) \end{array} \right] \\ \times \frac{\tilde{\theta}'(t+1)\tilde{Y}'(t+1)}{I'(t)} \\ + \frac{\tau_\omega|_{\vartheta'=1}(1-\tilde{\theta}'(t+1))\tilde{Y}'(t+1)}{N_o'(t+1)} \end{array} \right\}. \tag{4.3.40}$$

Illustrative Example Economy

Using (4.3.40), the central condition in the numerical example reads for the realization of the *first* state of $\tilde{Y}'(1)$, for instance,

$$c_o'^*(1)_1 = \frac{1}{1+\tau_C|_{\vartheta'=1}} \left\{ \begin{array}{l} \left[\begin{array}{l} i_y'(t-1)|_{\vartheta'=1} \frac{0.4 \times \$3,308.71}{125} \\ + (1-\tau_\omega|_{\vartheta'=1})(1-0.4) \frac{\$3,308.71}{125} \\ - (1+\tau_C|_{\vartheta'=1}) c_m'(t) \end{array} \right] \\ \times \frac{0.2 \times \$1,148.70}{\$200.00} \\ + \tau_\omega|_{\vartheta'=1} \frac{(1-0.2)\$1,148.70}{125} \end{array} \right\}.$$

In this equation the numerics of $\tau_\omega|_{\vartheta'=1}$ and $\tau_C|_{\vartheta'=1}$ have not been plugged in because their previous values have only been ad-hoc assumptions, which have not yet been derived properly.

The task is to establish them in such a way that above equation does not only hold for the first state's realization but also for the other ones. Furthermore, in order to replicate the complete markets' case perfectly it is obviously necessary that $c_m'(0)$ as well as $\tilde{c}_o'^*(1)_1$ match their unconstrained counterparts. This insight is crucial for the next section.

[16]This makes it also more similar to equation (50) of Merton [1983], as he does not address the accumulation of physical capital.

4.4 Replicating Tradability

Having established an alternative solution mechanism for the mimicking economy it can now be calibrated to replicate the benchmark model of Chap. 3. In other words, the parameters of the mimicking economy's public sector will be specified so that the resulting allocations of consumption and investments equal those of the economy with tradable human capital.

4.4.1 Replication Set-up

Since Sect. 4.2 established the mimicking economy independently from the original model, it is necessary to equate the respective variables in a first step. Secondly, the conditions for the replication have got to be set.

Assumptions

The mimicking economy shall be structurally identical to the benchmark model with the exception of tradability of human capital. In detail this assumptions implies that the population is the same in both economies, i.e.

$$N_y'(t) = N_y(t), \qquad N_m'(t) = N_m(t), \qquad N_o'(t) = N_o(t), \tag{4.4.1}$$

and parameters and stochastics equal as well:

$$\beta' = \beta, \quad \tilde{A}'(t) = \tilde{A}(t), \quad K_0' = K_0, \quad \vartheta' = \vartheta, \quad \tilde{\theta}'(t) = \tilde{\theta}(t). \tag{4.4.2}$$

The strong claim underlying (4.4.1) and (4.4.2) is that the deficiencies concerning the completeness of the capital market can be mitigated *perfectly* by the social planer's second-best solution. Nontradability of human capital does neither influence the decision about parenthood, the parametric structure of the economy nor its stochastic development. There is no difference for the agents between utility maximization with the human capital security and by participation in a pay-as-you-go pension system;[17] in particular their expectations regarding the sustainability of the public sector involvement is not distorted by skepticism regarding governmental involvement as political discretion has explicitly been eliminated in Sect. 4.2.2.

Replicating Conditions

In order to achieve the same equilibrium as in Chap. 3 the design of the fiscal system and of the PAYGO scheme must be so that it holds for all t

$$C_y'(t) \stackrel{!}{=} C_y(t), \quad C_m'(t) \stackrel{!}{=} C_m(t), \quad \text{and} \quad C_o'(t) \stackrel{!}{=} C_o(t) \tag{4.4.3}$$

[17]This is also reflected in the fact that agents' preferences in both economies are identical.

as well as

$$I'(t) \overset{!}{=} I(t). \tag{4.4.4}$$

Equation (4.4.3) states that the consumption of each generation must be the same in the unconstrained and in the replicating economy. Note that the identical demographic structure of (4.4.1) implies that also individual agents consume the same in the mimicking economy as in the benchmark model. To put it formally, if (4.4.3) is achieved it also holds that $c'_y(t) = c_y(t)$, $c'_m(t) = c_m(t)$, and $c'_o(t) = c_o(t)$. Equation (4.4.4) requires investments in physical capital to be the same in both economies. This condition and the same initial stock $K'_0 = K_0$ ensure the same accumulation of real capital. Of course, there is no condition on investments in human capital securities.

Implications

It is important to see that the conditions (4.4.3) and (4.4.4) are sufficient to perfectly replicate the benchmark model. Firstly, observe that (4.4.3) implies that aggregate consumption is identical:

$$C'(t) = C(t), \tag{4.4.5}$$

which follows by (3.5.2) and (4.3.16). With aggregate consumption and investments the same, demand for consumable units will also equal as by (3.2.3) and (4.2.5)

$$Y'(t) = C'(t) + I'(t) = C(t) + I(t) = Y(t)\,. \tag{4.4.6}$$

On the supply side, this is complemented by an identical development in the consumption and capital accumulation technologies. The capital adjustment process is the same for both economies because all periods' investments are the same based on (4.4.4) and parameters are identical due to (4.4.2). Hence, physical capital input into the consumption technologies is also identical. Since this technology's parameters do not differ either and since the simplicity of the labor market implies the same input of human labor, macroeconomic production is the same in the constraint and unconstrained case. With corresponding output and investments the investment-output ratio must be the same as can be seen from its definition:

$$\psi'(t) = \frac{I'(t)}{Y'(t)} = \frac{I(t)}{Y(t)} = \psi(t). \tag{4.4.7}$$

Since Chap. 3 revealed that $\psi(t)$ is a constant the replicating economy must have a constant investment-ouput ratio as well:

$$\Psi' = \psi'(t) = \psi(t) = \Psi\,. \tag{4.4.8}$$

To sum up, the allocations on the goods markets of the benchmark and mimicking economy are indistinguishable when the conditions (4.4.3) and (4.4.4)

are fulfilled.
Both economies' capital markets can obviously not be identical because the mimicking economy lacks the security tracking the value of human capital. Still, the financial market for physical capital securities is similar: with an identical process of physical capital accumulation firms in both economies have the same financing needs, i.e. the supply of securities will be equal. However, the structure of demand across generations in the mimicking economy is not required to match the unconstrained case. With the substantial changes in the budget constraints, savings are unlikely to be identical. There is also *no* need for $S'_y(t)$ equaling $S_y(t)$ or $S'_m(t)$ equaling $S_m(t)$, because agents care only about consumption and do now have access to other means for intertemporal transfers. For the underlying economic framework it does not matter either who owns the physical capital units as long as the total amount of investments allows to achieve the same accumulation of real capital as in the benchmark model.

4.4.2 Solving the Replication

The conditions and assumptions set up for the replication in the last section can now be used in the alternative characterization of the mimicking economy established in Sect. 4.3. This allows to derive an integrated condition on the parameters of the public sector, which can then be solved in order to obtain the desired specifications of τ_ω and τ_C.

Integrating the Conditions

To achieve this, one first focuses on the reduced form representations of the goods market equilibrium in the complete markets case from Sect. 3.6.3, namely (3.6.25), (3.6.26) and (3.6.27). These three equations have characterized each generation's consumption in the unconstrained case and must be maintained in the replication due to (4.4.3). Shifting them into the mimicking economy's prime notation and taking them to individual levels by dividing through cohort sizes reads

$$c'_y(t) = \frac{(1-\bar{\theta})\Psi'}{(\beta'^2+\beta'+1)\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]}\frac{\check{Y}'(t)}{N'_y(t)}, \tag{4.4.9}$$

$$c'_m(t) = \left[\frac{1}{\vartheta'}+\frac{1-\bar{\theta}'}{(\beta'^2+\beta'+1)\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]}\right]\frac{\Psi'}{\beta'}\frac{\check{Y}'(t)}{N'_m(t)} \tag{4.4.10}$$

$$\text{and}\quad c'_o(t) = \left[\frac{(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'}{\vartheta'}+\frac{1-\bar{\theta}'}{\beta'^2+\beta'+1}\right]\frac{\check{Y}'(t)}{N'_o(t)}. \tag{4.4.11}$$

These are the explicit conditions on agents' consumption for then replication.[18] It is then deducible from (4.4.8) that in the mimicking economy's

[18]Remember from Chap. 3 that $\check{Y}'(t)$ is just used to refer to the proper reduced form solutions.

equilibrium it must hold that

$$I'(t) = \Psi' \check{Y}'(t), \tag{4.4.12}$$

if the replicating conditions are to be fulfilled. The four equations (4.4.9) to (4.4.12) pin the mimicking economy's goods market allocations exactly to those obtained in the case of tradable human capital. As this is exactly what one wants to achieve, the conditions must also hold in the alternative characterization of the mimicking economy, i.e. in (4.3.36). Plugging them in gives a new central condition:

$$\left[\frac{(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'}{\vartheta'} + \frac{1-\bar{\theta}'}{\beta'^2+\beta'+1}\right]\frac{\tilde{\check{Y}}'(t+1)}{N_o'(t+1)} = \frac{1}{1+\tau_C}\times$$
$$\left\{\begin{array}{l} \left[\begin{array}{l} \frac{i_y'(t-1)}{N_m'(t)}\left(\theta'(t)\check{Y}'(t) + \frac{1-\vartheta'}{\vartheta'}\Psi'\check{Y}'(t)\right) + (1-\tau_\omega)(1-\theta'(t))\frac{\check{Y}'(t)}{N_m'(t)} \\ -(1+\tau_C)\left[\frac{1}{\vartheta'} + \frac{1-\bar{\theta}'}{(\beta'^2+\beta'+1)\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]}\right]\frac{\Psi'}{\beta'}\frac{\check{Y}'(t)}{N_m'(t)} \end{array}\right] \\ \times\frac{\tilde{\check{Y}}'(t+1)}{\Psi'\check{Y}'(t)}\left(\vartheta'\tilde{\theta}'(t+1) + (1-\vartheta')\Psi'\right) + \frac{\tau_\omega(1-\tilde{\theta}'(t+1))\tilde{\check{Y}}'(t+1)}{N_o'(t+1)} \end{array}\right\}. \tag{4.4.13}$$

Now note, that based on (4.2.3) one can cancel $N_m'(t)$ and $N_o'(t+1)$ from this equation. This is a remarkable step as it eliminates all dependencies on the demographic structure – which is exactly what has been required: a solution independent of the development in the population! Further laborious algebra – indicated in Appendix C.1.1 – allows to eliminate the fractions and to restate (4.4.13) finally as

$$\begin{aligned} 0 = {} & \theta'(t)\,\tilde{\theta}'(t+1)\,\beta'\vartheta'^2\Xi_2\left\{i_y'(t-1) - (1-\tau_\omega)\right\} \\ & +\theta'(t)\,\beta'\vartheta'(1-\vartheta')\Psi'\Xi_2\left\{i_y'(t-1) - (1-\tau_\omega)\right\} \\ & +\tilde{\theta}'(t+1)\,\vartheta'\left\{\begin{array}{l} i_y'(t-1)\beta'(1-\vartheta')\Psi'\Xi_2 + (1-\tau_\omega)\beta'\vartheta'\Xi_2 \\ -(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right] - \tau_\omega\Psi'\beta'\Xi_2 \end{array}\right\} \\ & +\Psi'\left\{\begin{array}{l} +i_y'(t-1)\beta'(1-\vartheta')^2\Psi'\Xi_2 \\ -(1+\tau_C)\beta'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right] \\ -(1+\tau_C)\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right](1-\vartheta') \\ +\beta'\vartheta'(1-\vartheta')\Xi_2 + \tau_\omega\beta'\vartheta'^2\Xi_2 \end{array}\right\}, \end{aligned} \tag{4.4.14}$$

where Ξ_2 is an auxiliary constant defined as $\Xi_2 \equiv \left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right](\beta'^2 + \beta'+1)$. Equation (4.4.14) is the integrated condition, which achieves the replication of first-best allocations. The alternative characterization developed before underlies this equation; however, intuition on individual terms is lost in course of the simplifications and only the characteristic parameters remain.

Public Sector Conditions

While (4.4.14) does not have the demographic variables any more, it does heavily depend on the stochastic realizations of elasticity parameter $\theta'(t)$ and $\tilde{\theta}'(t+1)$. Yet, the social planer intended to install a sustainable pay-as-you-go and fiscal system in the very simplest form, i.e. with *constant* contribution and tax rates as explained in Sect. 4.1.1. For this reason the parameters have been denoted τ_ω and τ_C rather than $\tau_\omega(t)$ and $\tau_C(t)$.
Similar to Merton [1983] one can now claim that (4.4.14) will hold for *any* sequential realizations of $\theta'(t)$ and $\tilde{\theta}'(t+1)$, if three conditions are met:

$$0 = i'_y(t-1) - (1-\tau_\omega), \tag{4.4.15}$$

$$0 = \left\{ \begin{array}{l} i'_y(t-1)\beta'(1-\vartheta')\Psi'\Xi_2 + (1-\tau_\omega)\beta'\vartheta'\Xi_2 \\ -(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right] - \tau_\omega\Psi'\beta'\Xi_2 \end{array} \right\}, \tag{4.4.16}$$

$$\text{and} \quad 0 = \left\{ \begin{array}{l} +\, i'_y(t-1)\beta'(1-\vartheta')^2\Psi'\Xi_2 \\ -(1+\tau_C)\beta'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] \\ -(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right](1-\vartheta') \\ +\, \beta'\vartheta'(1-\vartheta')\Xi_2 + \tau_\omega\beta'\vartheta'^2\Xi_2 \end{array} \right\}. \tag{4.4.17}$$

When (4.4.15) to (4.4.17) are fulfilled, each of the summands on the right hand side of (4.4.14) is zero; i.e. the condition holds and thus the replication of the complete markets' allocations will be achieved.
This represents a system of three equations with three unknowns, namely τ_C, τ_ω and $i'_y(t-1)$. Based on the first equation $i'_y(t-1)$ must obviously be independent of time as well, i.e. $i'_y(t-1) = i'_y(t) = i'_y$. Solving this system, as shown in Appendix C.1.2, gives the desired calibration.
The replicating contribution rate is given by

$$\tau_\omega = \frac{\left[(1-\vartheta')\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right]}{\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] + \Psi'(\Psi' + \vartheta' - 3\vartheta'\Psi' + 2\vartheta'^2\Psi')}. \tag{4.4.18}$$

The tax rate must be chosen according to

$$\tau_C = \frac{\left[\beta'\Psi'\Xi_2 - \beta'\vartheta'\Psi'\Xi_2 + \beta'\vartheta'\Xi_2\right] - \left[\Psi'\Xi_2 + \vartheta'\Psi' - \bar{\theta}'\vartheta'\Psi'\right]}{\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]}$$
$$- \frac{1}{\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]} \times$$
$$\frac{\beta'\Xi_2\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right]\left[(1-\vartheta')\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right]}{\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] + \Psi'(\Psi' + \vartheta' - 3\vartheta'\Psi' + 2\vartheta'^2\Psi')}. \tag{4.4.19}$$

And, for completeness, the young's fraction of investments is given by

$$i'_y = \frac{\Psi'\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] + \Psi'(\Psi' + \vartheta' - 3\vartheta'\Psi' + 2\vartheta'^2\Psi')}{\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] + \Psi'(\Psi' + \vartheta' - 3\vartheta'\Psi' + 2\vartheta'^2\Psi')}. \tag{4.4.20}$$

While τ_ω and τ_C can directly be determined by the social planer, the young's fraction of physical capital investment is realized indirectly: they receive benefits from the collected consumption tax, dedicate the replicating amount of this to consumption and save the rest. As real capital is the only form of savings in the constraint case this results in i'_y.

4.4.3 Analyzing the Solution

Having finally solved the model and established the appropriate rates in the public sector the properties of the solution and its numerical implications can be analyzed.

First Assessment

If the social planer implements the policies of a consumption tax and a pay-as-you-go public pension scheme with the rates as derived in the last section he can achieve a Pareto improvement in the sense that agents of all cohorts are able to realize exactly the same consumption paths as in the perfect markets' case. Because the condition for the replication, (4.4.3), required that the levels of consumption match *exactly* those of the unconstrained case, the second-best solution does not only ease the imperfection of nontradability but heals it completely!

In addition to this rather substantial efficiency gain, the proposed second-best solution has also the advantage that it is truly self-sustaining and does not need any intermediate decisions. In Sect. 4.2.2 it was required that the PAYGO as well as the tax system are balanced at all times; in this sense it is not necessary that any authority determines some economic or demographic adjustment factor to keep them sustainable over time. Instead, payments to retirees are completely financed by the contributions to the pension system and benefits to the young amount only to the aggregated revenues of the consumption tax. To emphasize it again: there are no deficits, in neither of the

public sector systems!
Yet, these advantages of the second-best solution come at a high price – measured in terms of total taxation. This can be seen when calculating the implied contribution and tax rates for the different values of the elasticity parameters in the consumption and capital adjustment technology. Figure 4.2 does this using Sect. 3's result for $\Psi' = \Psi$ and then (4.4.18) and (4.4.19) to derive the numerical values of τ_ω and τ_C dependent on ϑ' and $\bar{\theta}'$ for a subjective time preference parameter of $\beta' = 0.6$.
While τ_C seems to be in a reasonable range of up to about 25%, the required contribution rate τ_ω is relatively high with values never smaller than 60%. This means that employees can hardly keep a third of their wage! Of course this is implausible high compared to the typically observed contribution rates in real world PAYGO systems. Clearly, these numerical values are a major weakness of this analysis making the results rather unappealing and unusable for real-life politics. Yet, there are some good reasons for this mismatch of theory and practice – the simplicity of the model being an important first one and the implicit debts of real PAYGO systems a second one. While this will be left for the critical assessment in Chap. 5, the remainder of this section will focus on the sensitivities of the solution.

Properties of Solution

From eyeball-testing Fig. 4.2 one must ascertain that both rates are positive and smaller – or equal – than unity for all admissible values of the elasticity parameters. This is satisfying because a contribution or a tax rate exceeding 100% would have been as economically senseless as negative rates. Furthermore, τ_ω increases obviously with higher expected marginal productiveness of physical capital $\bar{\theta}'$, while τ_C decreases. The elasticity parameter in the adjustment cost function, ϑ', seems to have hardly any influence on the replication solution. Its influence on the rates is virtually undetectable in the three-dimensional figures. Yet, the effects of ϑ' are visible in the contour plots, depicted in Fig. 4.3. These illustrations show the same graphics as Fig. 4.2, but from such a perspective that the values of τ_ω and τ_C appear as contour lines in a truly two-dimensional image.
Even though the influence of ϑ' on τ_ω is small, it is clearly present in the most interesting region of $0.2 < \bar{\theta}' < 0.4$, which imply the typically observed distribution of income to physical capital and human labor. There, a high ϑ' implies slower ascending PAYGO contribution rates when $\bar{\theta}'$ is increased. For the tax rate the finding is different: for an increasing $\bar{\theta}'$ the decline in the replicating τ_C is more pronounced for medium values of the adjustment cost function's elasticity parameter and less when ϑ' approaches zero or one.

Intuition

It might seem surprising that the contribution rate to the PAYGO pension system increases with the expected marginal productiveness of physical capital, since labor's share of output – driven by $1 - \bar{\theta}'$ – is then smaller. Even

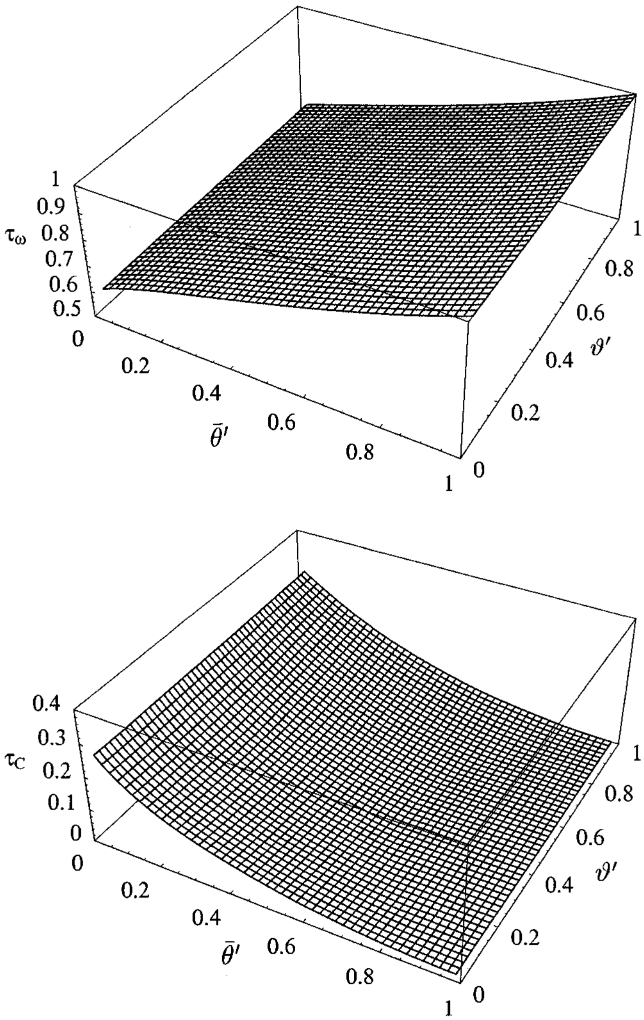

Fig. 4.2. Public sector parameters

Remarks: The plots depict the values of the contribution rate to the PAYGO scheme and the consumption tax rate that allow perfect replication of the complete markets' allocations for all admissible elasticity parameters, $0 < \vartheta' \leq 1$ and $0 \leq \bar{\theta}' \leq 1$, and a time preference parameter of $\beta' = 0.6$.

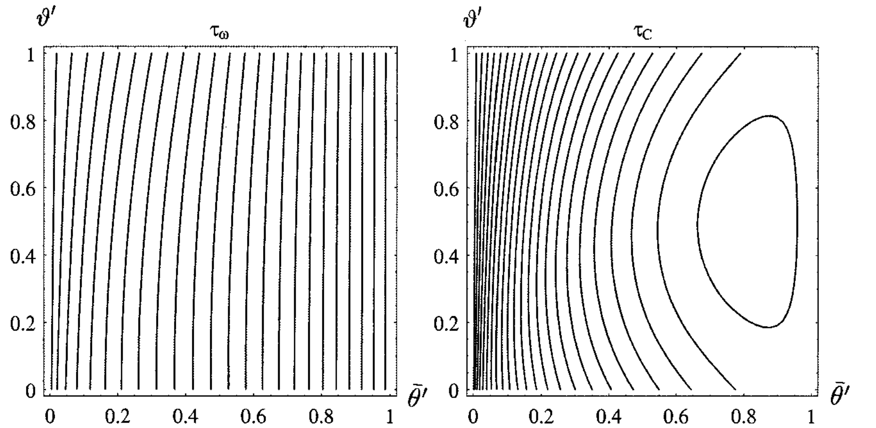

Fig. 4.3. Contour-plots of public sector parameters

Remarks: This figure illustrates the influence of the capital adjustment technology's elasticity parameter on the public sector parameters for $\beta' = 0.6$. The contour lines are ascending from left to right in the case of the PAYGO contribution rate and descending from left to right for the tax rate. While τ_ω is only slightly variant to ϑ' at lower levels of $\bar{\theta}'$, the elasticity parameter has more influence on τ_C.

the notion that the equilibrium wage rate falls with $\bar{\theta}'$ and that it requires thus a higher contribution rate to generate the same volume of PAYGO is not absolutely convincing, since the weight of human capital in the capital market of the underlying first-best model also declines. Yet another factor pushes in this direction. If $\bar{\theta}'$ is very high human labor is almost irrelevant for production and the return in the perfect capital market would be dominated by the return on physical capital. In such a situation participation in human capital and thus the PAYGO system is highly desirable for a middle-aged investor as it offers a vast diversification potential for him. In other words, at the extreme case of $\bar{\theta}' = 1$ middle-aged savers would only have the possibility to transfer consumption intertemporally by investing in real capital. In this case their marginal willingness to participate in the PAYGO system as a mean to trade human capital is extremely high – they would be willing to contribute their entire wage. As $\bar{\theta}'$ becomes smaller, physical capital loses its role as sole input factor reducing the diversification potential PAYGO offers and implying thus a lower τ_ω.
To understand the effect of the adjustment technology parameter one must refer to (4.2.6): with a high ϑ' new investments require less adjustment. Based on (4.2.8) this translates into the fact that fewer units of present consumption must be given up in order to realize the same value of physical capital stock. In this sense intertemporal consumption postponement by physical capital is less affected by the adjustment-costs making thus PAYGO less important.

However, this effect has only a quantitative effect on the contribution rate for the region of high ϑ' around $\bar{\theta}' \approx 0.3$, where the contour lines lean to the right. The reason for this might be that there the investment-output ratio is highest as seen in Fig. 3.3.
The decline in the replicating consumption tax rate when the expected marginal productiveness of physical capital increases can be explained by its role in the replication mechanism. When $\bar{\theta}'$ is low human labor is relatively important for production and receives thus a high share of output. In the benchmark model human capital is the risk-adjusted present value of this payoff and thus its current value is higher – ceteris paribus – the lower $\bar{\theta}'$. Since in the benchmark model the young generation's initial wealth is determined only by this endowment, they must proportionally be compensated when human capital is not tradable. As the revenues from consumption tax present the surrogate, these subsidies must be the higher the more valuable human capital would be in order not to make young agent worse off. Hence, the tax rate is higher when $\bar{\theta}'$ is low. The impact of the capital adjustment's elasticity parameter is presumably explained by the same argument. The various paths on which it enters (3.5.17) – as ϑ' , $1-\vartheta'$ and via the corresponding Ψ' – seem to counter each other somehow.
However, why ϑ has so little impact on either of the public sector parameters, stays unresolved to some extent. A reasonable conjecture to explain this difference to the analysis of Lim and Weil [2003] is that the long period length in the framework presented here substantially diminishes the adjustment cost effect. Note also, that the log-utility specification is very simple, whereas asset pricing approaches, like Jermann [1998], require habit formation in order for q-Theory to play a significant role.

4.4.4 Full Depreciation

To understand the impact of variations in agents' subjective time preference, it is again useful to restrict oneself to the case of fully depreciating physical capital. Since the results of this case can be derived by taking above limits or alternatively based on earlier results, a useful crosscheck for the findings is possible. Furthermore, excluding the accumulation of real capital is a first step towards embedding Merton's [1983] findings into this framework.

Limiting Case

Since the case of complete depreciation is characterized by setting the parameter $B' = 1$ and $\vartheta' = 1$ in the capital adjustment technology, the corresponding values of the public sector parameters are easily derivable. Taking the limits of (4.4.18), (4.4.19) and (4.4.20) yields the contribution and tax rate replicating the allocations of the complete markets' benchmark economy. According to Appendix C.2.1 these are

$$\tau_\omega|_{\vartheta'=1} = \frac{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}{1+\beta'+2\beta'^2\bar{\theta}'-\beta'^2\bar{\theta}'^2}, \tag{4.4.21}$$

$$\tau_C|_{\vartheta'=1} = \frac{\beta'^2(1-\bar{\theta}')^2}{1+\beta'+2\beta'^2\bar{\theta}'-\beta'^2\bar{\theta}'^2}, \tag{4.4.22}$$

$$\text{and} \quad i'_y|_{\vartheta'=1} = \frac{\beta'(1-\bar{\theta}'+\beta'\bar{\theta}'-\beta'\bar{\theta}'^2)}{1+\beta'+2\beta'^2\bar{\theta}'-\beta'^2\bar{\theta}'^2}. \tag{4.4.23}$$

Impact of Time Preference

For an analysis of the subjective time preference's impact on the replicating contribution and tax rate it is advantageous to focus on the above results. Figure 4.4 depicts the numerical values of $\tau_\omega|_{\vartheta'=1}$ and $\tau_C|_{\vartheta'=1}$ for all admissible $0 < \bar{\theta} < 1$: the replicating PAYGO contribution rate is strictly increasing in the time preference parameter; the consumption tax rate decreasing.[19]
To understand this it is insightful to investigate β''s impact on the investment-output ratio first. Based on (3.6.18) one sees that $\Psi'|_{\vartheta'=1} = \Psi|_{\vartheta=1}$ is increasing in the time preference parameter:

$$\begin{aligned}\frac{\partial\Psi'|_{\vartheta'=1}}{\partial\beta'} &= \frac{2\beta'\bar{\theta}'(1-\bar{\theta}')(1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}')-\beta'^2\bar{\theta}'(1-\bar{\theta}')(\bar{\theta}'+2\beta'\bar{\theta}')}{(1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}')^2}\\ &= \frac{\beta'\bar{\theta}'(1-\bar{\theta}')(2+\beta'\bar{\theta}')}{(1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}')^2} \geq 0,\end{aligned} \tag{4.4.24}$$

because $0 \leq \bar{\theta}' \leq 1$ and $0 \leq \beta' \leq 1$. Based on this an intuitive explanation of Fig. 4.4 is straightforward: when $\beta' = 1$ agents are indifferent to the timing of their consumption and save the most. This results in the highest investment-output ratio allowing for the largest output in the next period ceteris paribus. With a larger production labor's share of it in absolute terms and thus the wage level are both larger, requiring a smaller contribution rate in order to generate sufficient funds for the pay-as-you-go pension system. When time preference for present consumption increases, and thus β' decreases, the investment-output ratio declines commanding a growing contribution rate. At the extreme there would be no investment in physical capital – as seen in (3.6.19) – leaving PAYGO as the only mean for intertemporal consumption smoothing. In other words, for a high parameter of subjective time preference the intertemporal consumption postponement via funded savings works better.
The influence of the time preference parameter on the replicating consumption tax rate can be explained with a similar argument: with a high β', investments are relatively high compared to output and yield a large stock of physical capital in the future, which implies that human labor will be more productive.

[19] The influence of $\bar{\theta}'$ is the same as in the general case: the PAYGO contribution rate declines and the consumption tax rate increases with a growing marginal productiveness of physical capital.

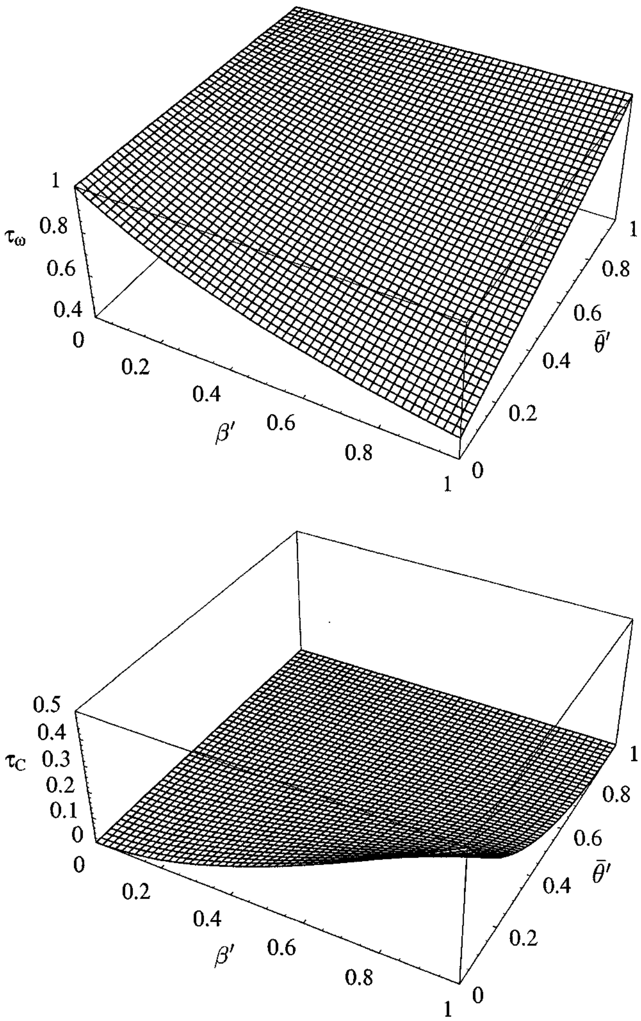

Fig. 4.4. Impact of time preference

Remarks: The illustration visualizes the impact of the subjective time preference parameter β' on the values of the PAYGO contribution rate and the consumption tax rate for the case of fully depreciating physical capital, $\vartheta' = 0$. At $\beta' = 1$ the plots also depict the case of Merton [1983].

Hence, the present value of future wages, i.e. the value of human capital, will be high enabling a large consumption of the young in the unconstrained economy, since their wealth budget is solely determined by this value. For the replication this means that the consumption tax rate must be higher in order to enable the young agents to realize the same level of consumption.

Alternative Derivation

In Sect. 4.3.5, the alternative budget constraint for fully depreciating real capital has already been established. Yet, there is another way to derive the results for this special case, which allows for a verification of (4.4.21) to (4.4.23). This approach nests on using the complete markets' results for cohort-specific consumption directly in the alternative budget constraint. Due to (4.4.3) and (4.4.4) it must hold in the mimicking economy for all t that

$$\begin{aligned} C'_y(t)|_{\vartheta'=1} &\overset{!}{=} C_y(t)|_{\vartheta=1}, & C'_m(t)|_{\vartheta'=1} &\overset{!}{=} C_m(t)|_{\vartheta=1}, \\ C'_o(t)|_{\vartheta'=1} &\overset{!}{=} C_o(t)|_{\vartheta=1}, & I'(t)|_{\vartheta'=1} &\overset{!}{=} I(t)|_{\vartheta=1}. \end{aligned} \tag{4.4.25}$$

Hence, the full depreciation's results of Sect. 3.6.4, (3.6.28) to (3.6.30), imply for the equilibrium allocations in the replication

$$C'_y(t)|_{\vartheta'=1} = \frac{1-\bar{\theta}'}{(\beta'^2+\beta'+1)\bar{\theta}'} I'(t), \tag{4.4.26}$$

$$C'_m(t)|_{\vartheta'=1} = \frac{\beta'(\beta'+1)\bar{\theta}'+1}{(\beta'^2+\beta'+1)\beta'\bar{\theta}'} I'(t), \tag{4.4.27}$$

$$\text{and} \quad C'_o(t)|_{\vartheta'=1} = \frac{\beta'(\beta'+1)\bar{\theta}'+1}{\beta'^2+\beta'+1} Y'(t). \tag{4.4.28}$$

Applying these allocations in the alternative budget constraint for a middle-aged agent, (4.3.40), reads

$$\begin{aligned} &\frac{\beta'(\beta'+1)\bar{\theta}'+1}{\beta'^2+\beta'+1}\frac{Y'(t+1)}{N'_o(t+1)} = \frac{1}{1+\tau_C|_{\vartheta'=1}} \\ &\times \left\{ \begin{aligned} &\left[\begin{aligned} &\frac{i'_y(t-1)|_{\vartheta'=1}}{N'_m(t)}\theta'(t)Y'(t) + (1-\tau_\omega|_{\vartheta'=1})(1-\theta'(t))\frac{Y'(t)}{N'_m(t)} \\ &-(1+\tau_C|_{\vartheta'=1})\frac{\beta'(\beta'+1)\bar{\theta}'+1}{(\beta'^2+\beta'+1)\beta'\bar{\theta}'}\frac{I'(t)}{N'_m(t)} \end{aligned} \right] \\ &\times \frac{\tilde{\theta}'(t+1)\tilde{Y}'(t+1)}{I'(t)} + \frac{\tau_\omega|_{\vartheta'=1}(1-\tilde{\theta}'(t+1))\tilde{Y}'(t+1)}{N'_o(t+1)} \end{aligned} \right\}. \end{aligned} \tag{4.4.29}$$

As shown in Appendix C.2.2 this will hold for *any* realizations of $\tilde{\theta}'(t)$ and $\tilde{\theta}'(t+1)$ if three conditions are met:

$$0 = i'_y(t-1)|_{\vartheta'=1} - (1 - \tau_\omega|_{\vartheta'=1}), \tag{4.4.30}$$

$$\begin{aligned} 0 = {} & (1 + \beta'\bar{\theta}' + \beta'^2\bar{\theta}') - \tau_\omega|_{\vartheta'=1}(1 + \beta'\bar{\theta}' + 2\beta'^2\bar{\theta}' - \beta'^2\bar{\theta}'^2) \\ & -\beta'(1 - \bar{\theta}')\frac{1 + \beta'\bar{\theta}' + \beta'^2\bar{\theta}'}{\beta'^2 + \beta' + 1} \\ & -\tau_C|_{\vartheta'=1}\beta'(1 - \bar{\theta}')\frac{1 + \beta'\bar{\theta}' + \beta'^2\bar{\theta}'}{\beta'^2 + \beta' + 1}, \end{aligned} \tag{4.4.31}$$

and $$0 = \tau_\omega|_{\vartheta'=1}(\beta'^2 + \beta' + 1) - (1 + \beta'\bar{\theta}' + \beta'^2\bar{\theta}')(1 + \tau_C|_{\vartheta'=1}). \tag{4.4.32}$$

Appendix C.2.2 also proofs that the solutions to this system of equations are again (4.4.21), (4.4.22) and (4.4.23). In other words, the alternative derivation of the special case with fully depreciating physical capital gives consistent results with taking the limits of the general solution.

4.4.5 Case of Merton [1983]

Having reduced the general results to the case of fully depreciating physical capital, one can now turn to the model of Merton [1983]. As mentioned in Sect. 2.3 he does not only abstract from real capital accumulation but also from time preference. However, human labor is supplied elastically in his case. Still, his results can easily be connected to this analysis.

First-best Allocations

To do so, note first that setting $\beta = \beta' = 1$ eliminates preference for present consumption in the agents' utility specification of (3.4.3) to (3.4.5) and (4.2.28) to (4.2.30). As explained in the last section agents become indifferent to the timing of their consumption when the subjective time preference parameter is one, because future utils do not need to be discounted.
While Merton [1983] abstracts from agents' time preference, he includes leisure in the utility specifications. Since this can only be relevant in an agent's working phase, utility for a young and a middle-aged are in his analysis extended to

$$\begin{aligned} U_y(t)|_{\text{Merton}} &= \ln c_y(t) + \mathbb{E}_t[\Gamma \times \ln l(t+1) + \ln c_m(t+1) + \ln c_o(t+2)] \\ &= \ln c_y(t) + \mathbb{E}_t[U_m(t+1)] \end{aligned} \tag{4.4.33}$$

and

$$\begin{aligned} U_m(t)|_{\text{Merton}} &= \Gamma \times \ln l(t+1) + \ln c_m(t) + \mathbb{E}_t \ln c_o(t+1)] \\ &= \Gamma \times \ln l(t+1) + \ln c_m(t) + \mathbb{E}_t[U_o(t+1)]. \end{aligned} \tag{4.4.34}$$

The only additional element is the fraction of a person's work period spent in leisure, $0 \leq l(t+1) \leq 1$, and a scaling factor $\Gamma \geq 0$ to include it in the utility function. Interesting insights on the design of PAYGO scheme can be obtained for elastic labor supply, i.e. for a nonzero Γ. These will be discussed

in Sect. 5.4.2. Yet, in order to link Merton's results to the presented framework one must focus on the case of inelastic labor supply as characterized by $\Gamma = 0$ and shown in Merton [1983, Sect. 12.3].
Since Merton abstracts also from real capital accumulation, his results for inelastic labor supply correspond to those of this framework when $\vartheta = \vartheta' = 1$, $B = B' = 1$ and $\beta = \beta' = 1$. The case without accumulation of physical capital, which is characterized by the first two conditions only, has been investigated extensively throughout Chaps. 3 and 4. It is most convenient to use this full depreciation case as a starting point for establishing the first-best solution and characterize its allocations.
Consequently, each cohort's consumption for the Merton case is derivable from the allocations of Sect. 3.6.4. Equations (3.6.28), (3.6.29) and (3.6.30) reduce for $\beta = 1$ to

$$C_y(t)|_{\text{Merton}} = \frac{1-\bar{\theta}}{(1^2+1+1)\bar{\theta}} I(t) = \frac{1-\bar{\theta}}{3\bar{\theta}} I(t), \tag{4.4.35}$$

$$C_m(t)|_{\text{Merton}} = \frac{1 \times (1+1)\bar{\theta}+1}{(1^2+1+1)1 \times \bar{\theta}} I(t) = \frac{1+2\bar{\theta}}{3\bar{\theta}} I(t), \tag{4.4.36}$$

$$\text{and} \quad C_o(t)|_{\text{Merton}} = \frac{1 \times (1+1)\bar{\theta}+1}{1^2+1+1} Y(t) = \frac{1+2\bar{\theta}}{3} Y(t). \tag{4.4.37}$$

These characterizations of the first-best allocation are given in Merton [1983] by equations (35), (37) and (38). Furthermore, the investment-output ratio corresponding to Merton's case is – based on (3.6.33) and (3.6.34) – given by

$$\Psi|_{\text{Merton}} = \frac{1^2 \times \bar{\theta} \times (1-\bar{\theta})}{1+1 \times \bar{\theta} + 1^2 \times \bar{\theta}} = \frac{\bar{\theta}(1-\bar{\theta})}{1+2\bar{\theta}}. \tag{4.4.38}$$

Consequently, investments are

$$I(t) = \Psi|_{\text{Merton}} \times Y(t) = \frac{\bar{\theta}(1-\bar{\theta})}{1+2\bar{\theta}} Y(t) \tag{4.4.39}$$

This corresponds to Merton's equation (39).

Replication with PAYGO

Scrutinizing on the case of inelastic labor supply, perfect replication of the first-best allocations requires that the conditions (4.4.3) and (4.4.4) hold analogously. Using (4.4.39), one can rewrite (4.4.36) and (4.4.37) in order to get the conditions on the allocations in the corresponding mimicking economy:

$$C'_m(t)|_{\text{Merton}} = \frac{1-\bar{\theta}'}{3} Y'(t), \tag{4.4.40}$$

$$C'_o(t)|_{\text{Merton}} = \frac{2\bar{\theta}'+1}{3} Y'(t). \tag{4.4.41}$$

Equations (4.4.39) to (4.4.41) are the same conditions Merton requires in his equations (47) to (49) for replication of the complete markets' equilibrium.[20] By integrating them into the corresponding alternative characterization of the middle-aged agents' budget constraint and solving the resulting equation system, Merton [1983] shows that perfect replication is achieved if

$$\tau_\omega|_{\text{Merton}} = \frac{1+2\bar{\theta}'}{2+2\bar{\theta}'-\bar{\theta}'^2}, \tag{4.4.42}$$

$$\tau_C|_{\text{Merton}} = \frac{(1-\bar{\theta}')^2}{2+2\bar{\theta}'-\bar{\theta}'^2}, \tag{4.4.43}$$

$$\text{and} \quad i'_y|_{\text{Merton}} = \frac{1-\bar{\theta}'^2}{2+2\bar{\theta}'-\bar{\theta}'^2}, \tag{4.4.44}$$

which are his equations (53a) to (53c). Since exactly these results could have also been derived by using $\beta' = 1$ in (4.4.21), (4.4.22) and (4.4.23), the case of Merton [1983] is perfectly embedded in this – more general – framework.

4.4.6 Illustrative Example Economy

Replication of the illustrative example economy is facilitated by the fact that the numerical example in this chapter has been chosen so that it perfectly matches its unconstrained counterpart of Chap. 3. In other words, the numerical assumptions in Sect. 4.2.5 have already fulfilled the replicating assumptions of (4.4.1) and (4.4.2) as far as possible. Therefore, the expected value for the elasticity parameter is again $\bar{\theta}' = 0.33 = \bar{\theta}$, as seen from Table 4.2. The only necessary extension is the notion that the subjective time discount factor is 0.6 and hence matching the complete markets case as well.

Allocations in Period 0

Consequently, one can immediately apply (4.4.21) to (4.4.23) in order to derive the parameters in the public sector and the investment fraction of the young. These can be calculated as

$$\tau_\omega = \frac{1+0.6\times 0.33+0.6^2\times 0.33}{1+0.6+2\times 0.6^2\times 0.33-0.6^2\times 0.33^2} \approx 73.2\%,$$

$$\tau_C = \frac{0.6^2(1-0.33)^2}{1+0.6+2\times 0.6^2\times 0.33-0.6^2\times 0.33^2} \approx 9.0\%$$

$$\text{and} \quad i'_y = \frac{0.6(1-0.33+0.6\times 0.33-0.6\times 0.33^2)}{1+0.6+2\times 0.6^2\times 0.33-0.6^2\times 0.33^2} \approx 26.8\%.$$

Now note that these figures match exactly the ad-hoc assumptions of Sect. 4.2.5. Therefore, one can use those results to derive the allocations in period 0.

[20] Note that Merton [1983] has a less rigorous approach concerning the notation and does hence not use primes for the structural parameters.

Starting with an old agent as beneficiary of the PAYGO system, it has been shown that such an agent's budget is \$15.53. Since he does not save, he dedicates this entire budget to consumption. Yet, he must pay the consumption tax of 9.0% on this, leaving him with a net consumption of

$$c_o'^*(0) = \frac{1}{1+9.0\%} \times \$15.53 \approx \$14.25.$$

Comparing this with his consumption in the unconstrained case, calculated in Sect. 3.4.6 as $c_o^*(0) = \$14.25$ results in the satisfying finding that he reaches exactly the same level of consumption and thus of utility.
For a middle-aged agent a wealth budget of \$7.09 from wages net of PAYGO contributions and previous savings was derived in Sect. 4.2.5. The fraction of physical capital investments corresponding to the middle-aged optimal savings is given through above figure and (4.3.25) as $i'_m = 1 - i'_y = 1 - 26.8\% = 73.2\%$. Furthermore, based on (4.4.8) the investment-output ratio of the mimicking economy matches the original 6.045%. This justifies Sect. 4.2.5's assumption of $I(0) = \$200$ since

$$I'(0) = \Psi' \times Y'(0) = 6.045\% \times \$3,308.71 \approx \$200.$$

Using these intermediate results, one can finally determine the real capital savings of the middle-aged generation as

$$S'_m(0) = i'_m \times I'(0) = 73.2\% \times \$200 \approx \$146.44$$

or for an individual agent as

$$s'^*_m(0) = \frac{S'_m(0)}{N_m(0)} = \frac{\$146.44}{125} \approx \$1.17.$$

Due to his budget constraint and the consumption tax this implies that a middle-agent's consumption is given by

$$c'^*_m(0) = \frac{1}{1+9.0\%} \times (\$7.09 - \$1.17) \approx \$5.43.$$

This does not only fulfill the optimality conditions (4.3.39) or (4.3.36) but is also the same amount as in the unconstrained case: Sect. 3.4.6 had established that $c^*_m(0) = \$5.43$. In other words, a middle-aged agent is as well off as with tradable human capital.[21]
Finally, one has for a young agent's real capital savings of

$$s'^*_y(0) = \frac{i'_y \times I'(0)}{N_y(0)} = \frac{26.8\% \times \$200}{100} \approx \$0.54$$

[21]Nota bene: the life-cycle consumption smoothing pattern does not manifest itself in the values \$5.43 and \$14.25, but only in the figures of the *same* cohort. Section 5.2.8 will address this further. The seemingly high consumption of the old reflects the no-bequest assumption.

allowing him together with the budget of \$2.79, as determined in Sect. 4.2.5, a consumption of

$$c_y'^*(0) = \frac{1}{1+9.0\%} \times (\$2.79 - \$0.54) \approx \$2.07.$$

Again this matches perfectly the level of consumption of such an agent in the unconstrained case![22] To sum up, all agents can achieve exactly the same levels of consumption as in the model with human capital securities. Furthermore, investments in real capital also match those of the unconstrained economy, so that the conditions for replication, (4.4.3) and (4.4.4), are satiesfied.

Consistency

Finally, remember from Sect. 4.2.5 that both the PAYGO and the fiscal system have a balanced budget. In period 0 the pension system collects contributions from the middle-aged of $\mathfrak{T}_o(0) = \$1,435.60$ and needs exactly this amount to finance total pension benefits to the old. For the subsidies to the young the fiscal system requires a budget of $\mathfrak{T}_y(0) = \$279.35$. To achieve this with the calculated tax rate of $\tau_C = 9.0\%$ an overall consumption of $C'(0) = \$3.108.71$ is required. But this is just the amount of consumption that will be realized, since consumption by all young, middle-aged and old agents adds up to

$$\begin{aligned} C'(0) &= \$2.07 \times 100 + \$5.43 \times 125 + \$14.25 \times 156 \\ &\approx \$207.17 + \$678.62 + \$2,222.91 \approx \$3.108.71. \end{aligned}$$

In other words, the assumption on aggregate consumption in Sect. 4.2.5 has already been the correct figure. Furthermore, output of consumable units has been calculated there as $Y'(0) = \$3.308.71$. Again the replication framework is consistent, as this figure coincides with demand from consumption and investment for these units. The goods market is in equilibrium since

$$Y'(0) = I'(0) + C'(0) = \$200.00 + \$3.108.71 = \$3.308.71.$$

Chapter 5 will give further insights on the consistency of the numerical example's replication.

4.5 Summary of Replication

In this chapter it has been shown that the first-best allocations of the general equilibrium model of Chap. 3 can be replicated for the case of nontradable human capital by a carefully designed pay-as-you-go public pension scheme and a general consumption tax. This section recapitulates the major steps, restates important formal results in Tab. 4.3 and reemphasizes major implications.

[22] Refer again to Sect. 3.4.6 where it has been calculated that $c_y^*(0) = \$2.07$.

Table 4.3. Summary for replication with PAYGO

		Achieved Replication		
Goods Market	$Y'(t) = Y(t)$			(4.4.6)
Investments	$I'(t) = I(t)$	$\Psi' = \Psi$		(4.4.4)
Consumption	$C'(t) = C(t)$			(4.4.5)
	$C'_y(t) = C_y(t)$	$C'_m(t) = C_m(t)$	$C'_o(t) = C_o(t)$	(4.4.3)
		Public Sector Parameters		
PAYGO Rate	$\tau_\omega = \frac{\left[(1-\vartheta')\Psi'+\vartheta'\right]\beta'\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]}{[2\Psi'-\vartheta'\Psi'+\vartheta']\beta'\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]+\Psi'(\Psi'+\vartheta'-3\vartheta'\Psi'+2\vartheta'^2\Psi')}$			(4.4.18)
Tax Rate	$\tau_C = \frac{\left[\beta'\Psi'\Xi_2-\beta'\vartheta'\Psi'\Xi_2+\beta'\vartheta'\Xi_2\right]-\left[\Psi'\Xi_2+\vartheta'\Psi'-\bar{\theta}'\vartheta'\Psi'\right]}{\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]} - \frac{1}{\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]} \frac{\beta'\Xi_2\left[2\Psi'-\vartheta'\Psi'+\vartheta'\right]\left[(1-\vartheta')\Psi'+\vartheta'\right]\beta'\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]}{[2\Psi'-\vartheta'\Psi'+\vartheta']\beta'\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]+\Psi'(\Psi'+\vartheta'-3\vartheta'\Psi'+2\vartheta'^2\Psi')}$			(4.4.19)
		Assumptions for Replication		
Population	$N'_y(t) = N_y(t)$	$N'_m(t) = N_m(t)$	$N'_o(t) = N_o(t)$	(4.4.1)
Parameters	$\beta' = \beta$	$\vartheta' = \vartheta$	$K'_0 = K_0$	(4.4.2)
Stochastics	$\tilde{A}'(t) = \tilde{A}(t)$	$\tilde{\theta}'(t) = \tilde{\theta}(t)$		(4.4.2)
		Mimicking Economy		
Public Sector	$\mathfrak{T}_o(t) = \mathfrak{t}_o(t) \times N'_o(t) = \tau_\omega \times \omega'(t) \times N'_m(t)$			(4.2.18)
	$\mathfrak{T}_y(t) = \mathfrak{t}_y(t) \times N'_y(t) = \tau_C \times C'(t)$			(4.2.22)
Capital Market	$\tilde{R}'_M(t+1) = \tilde{R}'_K(t+1) = \frac{\tilde{\theta}'(t+1)\tilde{Y}'(t+1)+\frac{1-\vartheta'}{\vartheta'}\tilde{I}'(t+1)}{K'(t+1)p'_K(t)}$			(4.2.12)
Real Sector	$Y'(t) = \tilde{A}'(t) \times K'(t)^{\tilde{\theta}'(t)} \times L'(t)^{1-\tilde{\theta}'(t)}$			(4.2.4)
	$K'(t+1) = B' \times I'(t)^{\vartheta'} \times K'(t)^{1-\vartheta'}$			(4.2.6)

Remarks: Chapter 4's replication of the first-best solution from Chap. 3 is summarized in this table. Nontradability of human labor can be mitigated by an appropriate pay-as-you-go pension system and a consumption tax. The constant Ξ_2 is defined as $\Xi_2 \equiv [(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'](\beta'^2 + \beta' + 1)$.

To start at the bottom of the table, the replication is based on a macroeconomic framework mimicking the complete markets' economy of Chap. 3. While the real sector of the overlapping generations model remains unchanged – comprising a technology for consumption goods and one for the accumulation of physical capital – major alterations have been observed on the capital market: since now there are no securities tracking the value of human capital, the capital market consists only of the financial market for physical capital and the market for the riskless asset. The latter is again in zero-net supply and thus the capital market's return is solely determined by the return on real capital. In order to realize the replication, the mimicking economy is supplemented by a public sector. This comprises a pay-as-you-go public pension scheme collecting contributions from the working middle-aged and paying the proceeds to the retired old. In addition to this, there is a general consumption tax financing a subsidy to the young. Both of the public sector systems maintain balanced budgets in order to achieve long-term sustainability.

It is then assumed that the development on the mimicking economy's real side is not affected by the imperfection on its capital market. The demographic structure as well as the parameters of the consumption and accumulation technology are absolutely identical to the unconstrained economy of Chap. 3. In particular the stochastic developments concerning the uncertainties in total factor productivity and in real capital's production elasticity remain unchanged. Consequently, the elimination of the hypothetical human capital securities and the addition of the public sector are the only changes compared to the earlier framework. However, even these two changes have a severe effect on solving the mimicking economy, since nontradability reintroduces nonadditivity in the agents' budget constraints. Yet, by referring to the aggregate levels the simplicity of the overlapping generations model allows to develop an alternative solution mechanism. With this implicit description of the economy it is possible to calibrate the PAYGO contribution rate and the consumption tax rate to the allocations of the unconstrained economy.

This achieves that agents of all generations can realize exactly those levels of consumption which they have achieved in the model with tradable human capital, and that total investments in real capital match those of the model with tradability. Since this allows for the same process of real capital accumulation on the mimicking economy's real sector and because the public sector systems maintain balanced budgets, this replication also holds dynamically. Due to the fact that only consumption determines the agents' utility rational this implies a second-best solution perfectly healing the imperfection of the mimicking economy's capital market.

To sum up, this chapter has extended the previous framework and shown that and how a PAYGO system and a consumption tax can replicate the diversification idea in the intertemporal consumption smoothing of the overlapping generations.

Part III

The Implications

5

Discussion and Assessment

It's far too late to escape the coming generational storm, but there's still time to get out of its direct path and avoid the full brunt of its impact.

Kotlikoff and Burns [2004]: "The Coming Generational Storm", p. 142

With the benchmark model and the replication of its allocations established this chapter analyzes the proposed framework and shows that it can cope with the requirements drawn in Chap. 2. The concept of national wealth is introduced in both settings because gross domestic product is not sufficient to account for all underlying economic relationships in an intertemporal perspective. While the explicit inclusion of technological progress does not yield any further diversification potential, the pay-as-you-go pension system can serve as a hedging instrument for uncertainty concerning the relative sizes of human labor's and physical capital's share of national income.

5.1 Overview

Chapter 4 has shown that and how the first-best model of Chap. 3 can be replicated with a pay-as-you-go pension scheme and a consumption tax. Before outlining this chapter's discussion, this section briefly highlights the major points of the replication and their consequences for the pension problem.

PAYGO and Pensionomics

The preceding chapters have developed a rather complex approach to the pension problem. The goal has been to find an economic rationale for sensibly combining funded and unfunded pension arrangements. As explained in Sect. 2.1, three dimensions of the problem had to be addressed: the existence of funded and unfunded pension provisions, the risks associated with each of them, and consistency in an aggregate macroeconomic perspective.

Since the budget of the PAYGO mechanism is centrally determined by the labor income of the working generation, the analytical approach to incorporate the portfolio idea has been an indirect one: instead of adding some PAYGO return directly to the standard portfolio problem, the analytically convenient trick of relaying on human capital as hypothetical kind of portfolio investments has been used. Because human capital is the capitalized value of future labor income, it does not only have the same connection to future wages as PAYGO, but also incorporates the corresponding risk adjustment in the discounting. In order to allow for some trading in the financial market for the assumed securities, there must be three generations in the overlapping generations framework as introduced in Sect. 3.1.1: a young generation that is able to issue these securities, a middle-aged one that has a motive to purchase them, and an old one claiming the payoffs on the previously acquired securities.

The feedback mechanism resulting from the pension savings of entire cohorts has been incorporated by an underlying macroeconomic framework. The technologies for the production of consumable output and for the accumulation of real capital, introduced in Sect. 3.2 and reestablished in Sect. 4.2, incorporate the important dimension of growth theory. This allows for a mutually consistent specification of future wages and of physical capital's rentals. Thereby the usage of adjustment costs is less to address q-Theory than to have a generalized version of the accumulation process: the specification of (3.2.7) embeds the special case of the Lucas [1978] economy as well as the version of Merton [1983]. The former model does *not* help for addressing the pension problem, even though it is the theoretical backbone for many financial models. In an endowment economy aggregate saving does not have any effect on the intertemporal process and the endowment stays completely exogenous. Given the expected scale of population changes as depicted in Fig. 1.1, this assumption can hardly be sustained. The Merton [1983]-version with completely depreciating physical capital, as the other extreme of (3.2.7), is a clear simplification of the framework, yet more intuitive for periods of 30 years. Therefore it has been addressed explicitly throughout the development of the replication.

The fact that the economic framework is in real terms and the existence of adjustment costs also allows for the notion of an asset-meltdown as feedback mechanism. The prices on the capital market have not been specified exogenously, but are determined by the saving motives of the living generations. Prices per se are not interesting; only the returns – and thus the terms-of-trade for intertemporal consumption transfer – they offer are relevant for the agents. Section 3.3 has reflected this view: future productive activity generates consumable output which is distributed to the agents as payoffs of both kind of securities. Expectations on these payoffs determine the current prices of the securities, so that the effect of a meltdown might be observed when a deteriorated outlook on future economic activity implies lower present prices than in the past. In essence, the young generation must replace the old one as partial owner of accumulated physical capital. And the young are only willing to pay according to the expected payoff the real capital stock generates.

The same is true for the human capital securities when they are marketable. When they are not marketable, there must be another mean to achieve the old's participation in labor income. This is the pay-as-you-go pension system as introduced in Sect. 4.2.2. As this allows only for an interaction between the middle-aged and the old generation an additional consumption tax system is required to replicate the second effect of tradable human capital: the young's trading of their human capital endowment in exchange for consumable output and participation in ownership of physical capital. The simplest and strictest form of sustainability is required for the public sector: no deficits and a constant contribution and tax rate.

When human capital is marketable, the agents' optimizing behavior can be solved straightforwardly for the equilibrium in the capital market as achieved in Sect. 3.5. Yet, when it is not tradable and the additional interaction of the agents with the public sector distorts the intertemporal budget constraint, the standard solution approach is not feasible. While for major parts of financial theory the simplifying exclusive focus on capital income might be sufficient, it would be deeply unjustified for the pension problem. Therefore, Sect. 4.3 established an alternative mechanism based on aggregate variables. Only the simple structure of the overlapping generations framework allowed this methodology: from the three living cohorts only two engage in savings. Hence, everything not owned by one of them must be owned by the other. Section 4.4 featured a second important analytical step toward replication. The observation that, with *constant* contribution and tax rates, the central equation (4.4.14) will hold for *any* realizations of $\theta'(t)$ and $\tilde{\theta}'(t+1)$, if the conditions of (4.4.15) to (4.4.17) are fulfilled, allowed to calibrate the public sector parameters *independently* of the stochastics. Yet, this is not a necessary condition and generalizations with dynamic rules for $\tau_\omega(t)$ and $\tau_C(t)$ may well be found. In this sense the replication limits itself to the most simplest form. Still, it is the final step of the indirect implementation of the portfolio idea: the calibrated PAYGO system fulfills the role of tradable human capital by giving the old a stake in labor's share of output. Funded and unfunded pension arrangements are clearly incorporated in the replication. The first-best benchmark model is just an extremely useful auxiliary construct for establishing the risk-considerations in the portfolio. And the desired feedback mechanisms are deeply rooted in the macroeconomic framework. The rigorous relation of the PAYGO mechanism to and the derivation of the return specification from it have shown how closely the financial and the public sector are tied to the economic framework. There will be no answer to the pension problem without addressing the underlying economics. Therefore – in an era of global aging – Pensionomics should be of increasing importance.

Outline

The remainder of this chapter will look at the above discussed aspects at more detail. Section 5.2 shows formally that the framework can cope with the requirements established in Chap. 2: the pricing of the assets reflects finance's standard view of prices as risk-adjusted discounted future payoffs. The portfolio idea is equally well incorporated in the model. It reflects the consumptive aspect of savings as well as the resulting accumulation of physical capital.

Yet, the standard national income accounting framework must be adjusted in order to thoroughly include the human capital securities. In order to properly reflect the intertemporal perspective it must be supplemented by the concept of national wealth. Furthermore, it is shown in this section that the return on PAYGO serves as a surrogate for the human capital securities when they are not marketable. Growth theoretic aspects incorporated in the macroeconomic production and accumulation technologies have been a cornerstone of the entire framework and are thus clearly present. However, technological progress as an important issue of growth theory has only been addressed imprecisely. Therefore, in Sect. 5.3 a brief continuous time version of a stochastic growth model incorporates Harrod-neutral progress and shows that this does not lead to additional diversification potential. Instead, only the uncertainty of the production elasticity determines the relative sizes of labor's and physical capital's share of income and is thus the key driver for diversification. Section 5.4 extends this view and focuses on the diversifiable and undiversifiable risk components. Based on this understanding the role of earnings points and the effect of wage indexation of pension benefits, which are both typically present in existing pay-as-you-go schemes, is critically analyzed. Finally, it is argued that current PAYGO systems are deeply unsustainable and that this is the reason why they can operate with a substantially lower contribution rate than implied by this framework. The summary of Sect. 5.5 reemphasizes this view of risk sharing.

5.2 Analyzing the Framework

Given the rather complex formal replication model it is helpful to crosscheck the results and see how the different pieces fit together. According to Chap. 2, the framework is supposed to cope with standard findings of economic science: the simultaneous consideration of funded pension savings and PAYGO participation should neither contradict finance nor economic theory concerning portfolio formation and long-run economic growth. Since the replication rests on the same economic framework, either of the model's versions can be used to address these issues unless the capital market incompleteness plays an explicit role. Obviously, the more tractable complete market case is more convenient and will hence be in the primarily focus. The replication is only considered when the particularities of the PAYGO system are addressed.

5.2.1 Asset Pricing and Allocation

Starting with financial theory, one of its standard paradigms is that prices are net present values of future payoffs with risk-adjusted discounting. The model's results on security prices should reflect this asset pricing view in the classical manner as well as in the framework of stochastic discount factors.

Real Capital Security

The interpretation of prices as present values has already been given when describing the real capital security in Sect. 3.3.1. In particular, (3.3.3) stated that the market value of the real capital securities, $K(t+1)p_K(t)$, is given by the present value of future rentals of physical capital, $\tilde{\theta}(t+1)\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)$, discounted at the risk-corresponding rate of $\tilde{R}_K(t+1)$:

$$K(t+1)p_K(t) = \mathbb{E}_t\left[\frac{\tilde{\theta}(t+1)\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}{\tilde{R}_K(t+1)}\right].$$

To show how the general equilibrium solution fulfills this condition take the finding for the return on real capital of (3.5.23), namely

$$\tilde{R}_K(t+1) = \left[(1-\vartheta)\tilde{\psi}(t+1) + \vartheta\tilde{\theta}(t+1)\right]\frac{\tilde{Y}(t+1)}{I(t)},$$

make it an explicit reduced form by replacing $\tilde{\psi}(t+1)$ with Ψ according to (3.6.15) and insert it into the present value statement from before:[1]

$$K(t+1)p_K(t) = \mathbb{E}_t\left[\frac{\tilde{\theta}(t+1)\tilde{Y}(t+1) + \frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}{[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)]\frac{\tilde{Y}(t+1)}{I(t)}}\right].$$

Bringing $K(t+1)$ to the right hand side, taking known terms out of the expectations operator and applying the investment-output ratio gives

$$p_K(t) = \mathbb{E}_t\left[\frac{\vartheta\tilde{\theta}(t+1) + (1-\vartheta)\Psi}{(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)}\right]\frac{1}{\vartheta}\frac{I(t)}{K(t+1)}.$$

As nominator and denominator are the same the expectation of the fraction is one; hence

$$p_K(t) = \frac{1}{\vartheta}\frac{I(t)}{K(t+1)} = \frac{1}{\vartheta}\frac{I(t)}{B\,I(t)^{\vartheta}K(t)^{1-\vartheta}} = \frac{1}{B\vartheta}\left(\frac{I(t)}{K(t)}\right)^{1-\vartheta}, \quad (5.2.1)$$

where (3.2.7) has been applied for $K(t+1)$. Nota bene, that this is exactly the same statement for the price of a unit of physical capital as the measure derived in Sect. 3.2.2 based on opportunity costs. In other words (5.2.1) equals (3.2.11), which proofs the internal consistency of the framework.[2]
Alternatively, one could use asset pricing's standard concept of stochastic discount factors and the Euler condition for the expected return on the physical capital security. According to (3.5.28) it must hold that

[1]As mentioned in Sect. 3.6.3, one actually needs to restate everything with the reduced form representations.

[2]Since the real capital securities are also present in the constraint version, this holds also for this case.

$$1 = \mathbb{E}_t \left[\tilde{M}_y(t+1) \times \tilde{R}_K(t+1) \right].$$

Inserting the general equilibrium solutions (3.5.29) for $\tilde{M}_y(t+1)$ and again (3.5.23) for $\tilde{R}_K(t+1)$ reads[3]

$$\mathbb{E}_t \left[\tilde{M}_y(t+1) \, \tilde{R}_K(t+1) \right]$$
$$= \mathbb{E}_t \left[\frac{1}{(1-\vartheta)\Psi + \vartheta\bar{\theta}} \frac{(1-\vartheta)\Psi + \vartheta}{(1-\vartheta)\Psi + \vartheta} \frac{I(t)}{\tilde{Y}(t+1)} \frac{[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)]\tilde{Y}(t+1)}{I(t)} \right].$$

This simplifies to

$$= \mathbb{E}_t \left[\frac{(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)}{(1-\vartheta)\Psi + \vartheta\bar{\theta}} \right] = \frac{(1-\vartheta)\Psi + \vartheta\,\mathbb{E}_t\left[\tilde{\theta}(t+1)\right]}{(1-\vartheta)\Psi + \vartheta\bar{\theta}} = 1, \qquad (5.2.2)$$

because $\mathbb{E}_t[\tilde{\theta}(t+1)] = \bar{\theta}$ according to the assumptions in Sect. 3.2.1. This exercise proofs that the general equilibrium's solutions for the SDFs and the return on the real capital security fulfill (3.5.28). Thus the model is well nested in the asset pricing framework.

Human Capital Security

The same crosschecks can be applied for the securities of the human capital kind. The present value interpretation of (3.3.9) implies that the current price of such a security, $p_H(t)$, represents the expected value of payoff from wages, $\tilde{\omega}(t+1)$, discounted at the rate corresponding to its risk characteristics, $\tilde{R}_H(t+1)$. Using (3.3.7) for the equilibrium wage rate one has

$$p_H(t) = \mathbb{E}_t \left[\frac{(1-\tilde{\theta}(t+1))\frac{\tilde{Y}(t+1)}{\tilde{N}_m(t+1)}}{\tilde{R}_H(t+1)} \right]. \qquad (5.2.3)$$

Plugging the general equilibrium result of (3.5.24), namely

$$\tilde{R}_H(t+1) = (1-\tilde{\theta}(t+1)) \frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{1-\bar{\theta}} \frac{\tilde{Y}(t+1)}{I(t)},$$

into (5.2.3) and simplifying using the demography assumption $\tilde{N}_m(t+1) = N_y(t)$ gives

$$p_H(t) = \mathbb{E}_t \left[\frac{(1-\tilde{\theta}(t+1))\frac{\tilde{Y}(t+1)}{\tilde{N}_m(t+1)}}{(1-\tilde{\theta}(t+1))\frac{(1-\vartheta)\Psi+\vartheta\bar{\theta}}{1-\bar{\theta}}\frac{\tilde{Y}(t+1)}{I(t)}} \right] = \frac{1-\bar{\theta}}{(1-\vartheta)\Psi + \vartheta\bar{\theta}} \frac{I(t)}{N_y(t)} \qquad (5.2.4)$$

Taking now $N_y(t)$ onto the left hand side of (5.2.4), the equation becomes the expression for the market capitalization of human capital in equilibrium, (3.5.17), and is hence just another illustration of the framework's consistency. Again this could have also been achieved in the SDF framework.[4]

[3] Again Ψ replaces $\bar{\psi}$ as well as $\tilde{\psi}(t+1)$ in the general equilibrium solution!
[4] See Appendix D.1.

Riskless Security

Finally, the solution for the shadow riskless security also meets the conditions imposed on it.[5] First, note that the general equilibrium (3.6.15) allows to simplify the condition (3.5.26) to:

$$\begin{aligned} R(t) &= \frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{(1-\vartheta)\Psi + \vartheta} \frac{1}{I(t)} \left(\mathbb{E}_t \left[\frac{1}{[(1-\vartheta)\Psi + \vartheta]\, \tilde{Y}(t+1)} \right] \right)^{-1} \\ &= \frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{I(t)} \left(\mathbb{E}_t \left[\frac{1}{\tilde{Y}(t+1)} \right] \right)^{-1} . \end{aligned} \tag{5.2.5}$$

For the riskless security the Euler equation of (3.5.28) can be reformulated as

$$\frac{1}{R(t)} = \mathbb{E}_t \left[\tilde{M}_y(t+1) \right] . \tag{5.2.6}$$

As well known, the inverse of the riskless return equals the expectation of the stochastic discount factor. To crosscheck this take the expectations of the general equilibrium version of (3.5.29) and simplify

$$\begin{aligned} \mathbb{E}_t \left[\tilde{M}_y(t+1) \right] &= \mathbb{E}_t \left[\frac{1}{(1-\vartheta)\Psi + \vartheta\bar{\theta}} \frac{(1-\vartheta)\Psi + \vartheta}{(1-\vartheta)\Psi + \vartheta} \frac{I(t)}{\tilde{Y}(t+1)} \right] \\ &= \frac{I(t)}{(1-\vartheta)\Psi + \vartheta\bar{\theta}} \mathbb{E}_t \left[\frac{1}{\tilde{Y}(t+1)} \right] \end{aligned} \tag{5.2.7}$$

Since (5.2.7) is the inverse of (5.2.5) this security fits well into the standard asset pricing framework, too.

Myopia

Turning to the aspects of asset allocation in the agents' portfolio, the policies developed in Sect. 3.4.3 clearly reflect the very standard portfolio approach from finance. While the particular features of the concept here are quite simple compared to models that scrutinize on the portfolio formation problem only, the asset allocation still reflects the essential features of risk diversification and return maximization. This is driven by the underlying assumptions on utility reflecting the two classical items of investment behavior for financial research: non-satiation and risk aversion. However, the framework exhibits myopia in the portfolio rules as defined in Ingersoll [1987, p. 257] or Merton [1990, p. 170]. That is, the asset allocation does not depend on any state variable, on the agent's wealth or on the planning horizon.[6] The reason for this is

[5] As there have not been made any assumptions on the price of a riskless bond, a present value statement cannot be checked. Yet, Appendix D.1 provides an interesting extension in this direction.

[6] One might argue that the valuation of the young's endowment with human capital is driven by the development in the economy's real side. However, even this indirect link to the state variables does not affect the portfolio policy but only the size of the portfolio.

Sect. 3.4.1's assumption of log-utility. This form for the utility function eliminates any hedging demand with respect to the economy's state of nature – here characterized by $\tilde{A}(t)$ and $\tilde{\theta}(t)$ – as shown in the classic results of Samuelson [1969] and Merton [1969*b*, 1971] and nicely summarized in Campbell and Viceira [2001*a*, Chap. 2]. While this simplification is a clear drawback of the framework presented here, it also made it possible to establish equilibrium in the capital market: Remember from Sect. 3.5.2 that only the identity of the young's and middle-aged's portfolio rules, (3.4.27) and (3.4.18), allowed to identify the market portfolio's weights being identical with these agent-specific portfolio weights. If there was some state contingency, this aggregation would be seriously impeded. Moreover, the desirable generalization to power utility would also require additional assumptions. Since Samuelson [1969] and Merton [1969*b*, 1971] have also shown that portfolio choice is still myopic for power utility and independent and identically distributed returns, one must prevent human and real capital's return to exhibit this characteristic. Given that the i.i.d.-assumption is rather reasonable for $\tilde{\theta}(t)$,[7] the state variable for total factor productivity is the only way to prevent the returns from exhibiting the same distributional characteristic. Given the aim of consistency with growth theory, it is a remarkable task to establish an analytically convenient but as well theoretically founded assumption for $\tilde{A}(t)$; even more so as the impact of aging populations on progress is unclear.
Another justification for the simple case with log-utility is the fact that the simple structure of the overlapping generations model has leeway only for a single rebalancing of the portfolio: with only three periods to live an agent would have the opportunity to adjust a previously chosen portfolio only when he is middle-aged. Increasing the number of rebalancing implies splitting the life-cycle in more periods and hence adding additional cohorts to the model. While this might be realizable for the complete markets' case, the resulting dynamic portfolio rule will present a major challenge for the replication with PAYGO. It is very unlikely that such a simple pay-as-you-go scheme with constant contribution rates as presented in Chap. 4 will still do the job. Finally, note that the asset allocation aspect is not driving the results of Constantinides et al. [2002] either. Because human capital is not tradable in their framework, real capital is a diversifying investment with and without dynamic portfolio allocations.

Consumption Smoothing

The simple the utility assumption might be it still achieves the most important pattern of financial asset-pricing models: smoothing of consumption. As explained in Sect. 2.2.1, this is a key feature of economic theory and the model clearly captures it because *any* form of income enters wealth as defined in (4.2.28) to (4.2.30) and is thus available for consumption. A windfall income,

[7] See Sect. 3.2.1 and note that an assumption has not been made on the distribution of $\tilde{A}(t)$.

regardless whether it is due to a high return on the previous investments, due to a high wage rate or due to high transfer income, leads to a higher wealth. Based on the saving rule this windfall is not entirely consumed, but spread across all during which the agents lives.

Yet, note an important difference to the manifold studies of consumption-smoothing and saving throughout a business cycle, like for instance Fama and French [1989]. These contributions are driven by a demand-side perspective: in bad times income is low and expected returns must be high to induce agents' substitution from consumption into investments for a *fixed* number of securities. In contrast to this, the model here, with its neoclassical production framework, is a supply-side driven long-run view. Hence, the number of real capital securities is not fixed but endogenous as described in Sects. 3.3.1 and 3.3.4: whenever agents give up present consumption and allocate savings to physical capital, the firms' managers receive more financing and can then acquire the unconsumed units on the goods market. Thereby they issue additional real capital securities. Therefore prices on the capital market are not driven up or depressed purely because of the present business conditions, but reflect rationally the long-run perspective of the economy given the present investment behavior of its agents. Of course, even in this long term perspective agents require compensation for uncertainty of their investments.

5.2.2 National Income

These two perspectives of savings – postponing consumption by savings and increasing real capital via investments – brings one to the macroeconomic aspects of the model. Addressing both of them consistently is a key point of the framework. By the explicit formulation of the capital adjustment technology, the accumulative effect has been thoroughly incorporated and extensively studied. Concerning the consumptive effect, on the other hand, Sect. 3.6.1 has revealed a major deviation from standard national income accounting.

Savings versus Investments

Because the introduction of the human capital securities represents a non-standard extension, the definitions of aggregate savings and investments in physical capital applied in this framework do *not* imply that both are equal; i.e. $I(t) \neq S(t)$! This can directly be seen from the general equilibrium form of (3.6.3):

$$S(t) = \frac{1}{\vartheta}\frac{(1-\vartheta)\Psi + \vartheta}{(1-\vartheta)\Psi + \vartheta\theta} I(t).$$

The reason for this is twofold: First, $I(t)$ refers only to physical savings, i.e. real capital investments, while $S(t)$ does also include allocations to the human capital securities. Aggregate savings must equal the total market value of *both*, physical and human capital, as in equilibrium all securities are held.

Second, the convex adjustment technology distorts the equality by the factor $1/\vartheta \geq 1$ because saved consumption units do not become installed capital on a one-by-one basis, and physical capital is valued higher then the underlying unconsumed units. This is another reflection of Sect. 3.3.1's finding, that with convex adjustment costs installed capital has a value per se.
Hence, by accounting for the first distinction and by addressing the convex adjustment costs the equality of savings and investment can be restored. To do so, observe that savings allocated to real capital, $S_K(t)$, are given by the fraction of total savings dedicated to physical capital, i.e. total savings times the market portfolio's weight of physical capital:

$$S_K(t) = S(t) \times \alpha_K^M(t). \tag{5.2.8}$$

Using the equilibrium results of (3.5.20) and (3.6.3) allows to reduce this to

$$S_K(t) = \frac{(1-\vartheta)\Psi + \vartheta}{(1-\vartheta)\vartheta\Psi + \vartheta^2\bar{\theta}} I(t) \times \frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{(1-\vartheta)\bar{\Psi} + \vartheta} = \frac{1}{\vartheta} I(t). \tag{5.2.9}$$

As the mimicking economy lacks the human capital securities the actual distinction in savings becomes unnecessary because (4.3.18) and (4.3.19) implied $S'_K(t) = S'(t)$. Still, (5.2.9) holds equivalently for the replication with PAYGO as can be seen from (4.3.20) to (4.3.22), since this is the effect of the adjustment costs, which are also present in the mimicking economy.
To understand their influence, note first that the increasing effect of investments on the physical capital stock is measured by $S_K(t)$, while $I(t)$ alone captures only the non-consumptive effect of savings in the goods market. In this sense savings allocated to real capital – but not $I(t)$ – capture the effect of the adjustment costs. Because next period's physical capital stock must be purchased each period on the output market, agents must be willing to forgo consumption of $I(t)$ consumable units. Yet, because savings include the additional benefit in form of the physical capital's value per se, the opportunity costs of this forgone consumption is smaller; in other words agents value the corresponding savings $S_K(t)$ higher than $I(t)$. Finally, note that in (3.2.12) – or equivalently in (4.2.8) – a valuation of next period's physical capital stock has already been established. The value has been found to be $1/\vartheta I(t)$, which is consistent with (5.2.9).
This becomes even clearer when the firms' financing decision is addressed explicitly. In order to buy consumable units on the output market, firms issue real capital securities. Yet, they are sold to investors on the financial market at a price higher than the mere opportunity costs of forgone consumption because the real capital – once installed – has got the value per se and managers as well as investors understand this fully. Hence, equilibrium on the market for real capital securities is achieved when the supply, $K(t+1)p_K(t)$, equals the demand by investors $S_K(t)$. As the former equals $1/\vartheta \times I(t)$ according to (3.2.12), one has established again

$$S_K(t) = K(t+1)p_K(t) = \frac{1}{\vartheta} I(t).$$

In the interpretation of q-Theory, investors are willing to buy fractional ownership of firms at a price higher than acquiring the duplicating consumption units on the output market because the installation of them – the formation of a firm – is of value as well.

Gross Domestic Product

It is insightful to connect these results on savings and investment with the findings for the gross domestic product of Sect. 3.3.1. It was shown there that

$$Y^{GDP}(t) = C(t) + \frac{1}{\vartheta}I(t).$$

Rearranging this and using (5.2.9) allows to write

$$Y^{GDP}(t) - C(t) = \frac{1}{\vartheta}I(t) = S_K(t). \tag{5.2.10}$$

Savings allocated to real capital equal the part of gross domestic product that is not consumed. As total savings can only be realized in form of allocations to real or human capital – and the market portfolio weights add up to one – it must hold that

$$S_H(t) = S(t) - S_K(t). \tag{5.2.11}$$

Yet, unlike the supply of real capital securities, which is endogenously determined by the financing decision of the firm, their human capital counterparts – if tradable – are added to the economy according to the births of young agents in the population. Note that the finding for gross domestic product, (5.2.10), holds correspondingly for the mimicking economy, too, since the incompleteness on the capital market does not interfere with this aspect of the real sector.[8] Yet, the distinction of (5.2.11) becomes trivial again, since $S'_H(t) = 0$.

Having just shown that gross domestic product includes only the ***physical*** or funded component of savings the question arises, where the fraction allocated to the human capital securities is accounted for. To address this in the economic accounting framework one must refer to the concept of national wealth.

5.2.3 National Wealth

The integration of the economy's endowment with the human capital of its young generation can only be achieved with the concept of wealth as analogously introduced in the last chapters. This dimension is most clearly introduced in the benchmark model but will allow important insights on the consequences of PAYGO. Before turning to these in Sect. 5.2.4, it is easier to focus on the complete markets' case only.

[8]For this reason it has been possible to establish $Y'^{GDP}(t)$ in (4.2.9) absolutely identical to $Y^{GDP}(t)$ in Sect. 3.3.1.

Aggregation

The assumptions in Sect. 3.4.2 have been made on the level of the individual agent. It was stated in (3.4.6) that each a-aged agent can use his wealth for either consumption or savings – regardless of the allocation:

$$w_a(t) = c_a(t) + s_a(t)$$

Section 3.5.1 has taken this definition to the cohort level and established in (3.5.3) and (3.5.5) that

$$W_y(t) = C_y(t) + S_y(t), \quad W_m(t) = C_m(t) + S_m(t), \quad \text{and} \quad W_o(t) = C_o(t),$$

where one must remember that the old cohort does not save. Now use (3.5.2) to get the aggregated equivalent of (3.4.6), namely

$$\begin{aligned} W(t) &= W_y(t) + W_m(t) + W_o(t) = C_y(t) + S_y(t) + C_m(t) + S_m(t) + C_o(t) \\ &= C(t) + S(t). \end{aligned} \tag{5.2.12}$$

The economy's national wealth,[9] $W(t)$, can be used for either aggregated consumption, $C(t)$, or aggregated savings, $S(t)$. Nota bene: according to (5.2.12) national wealth includes *all* forms of savings. In contrast to this (5.2.10) states that only savings allocated to real capital enter the concept of gross domestic product,

$$Y^{GDP}(t) = C(t) + S_K(t). \tag{5.2.13}$$

Combining these equations by usage of (5.2.11) relates national wealth to gross domestic product

$$W(t) = C(t) + S(t) = C(t) + S_K(t) + S_H(t) = Y^{GDP} + S_H(t). \tag{5.2.14}$$

In other words, national wealth exceeds gross domestic product by the amount of savings allocated to human capital securities.

Market for Human Capital Securities

According to Sect. 3.3.4 the supply of these securities is determined by the young cohort's size. Since each young agent enters the economy with a wealth endowment of one security of the human capital kind, equilibrium in this financial market requires

$$S_H(t) = p_H(t) N_y(t). \tag{5.2.15}$$

But combining this with the definition of the young cohort's wealth, (3.5.1), means

[9]The reference to *national* rather than total or aggregated wealth follows Merton [1983].

$$S_H(t) = W_y(t). \tag{5.2.16}$$

Aggregate savings in human capital – i.e. not only the portfolio allocation of the middle-aged – corresponds to the wealth of the young cohort. To understand the implication of this, it is useful to hypothetically decompose the behavior of the young: with the financial market for human capital in place they can first sell their future labor income *completely*, decide on consumption and savings and then repurchase human capital securities in an amount corresponding to the desired portfolio allocation. In this sense, they borrow against the *entire* capitalized value of their human capital. Equation (5.2.16) states that the valuation in this first and complete sale of securities is determined by aggregate demand for human capital savings, in other words by the amount middle-aged and young agents themselves jointly allocate to it.
Combining (5.2.16) and (5.2.14) establishes $W(t)$ as

$$W(t) = Y^{GDP} + W_y(t). \tag{5.2.17}$$

The concepts of national wealth and gross domestic product differ by the aggregated wealth of the young generation. This is not a surprise, given the fact that it was assumed that the young agents are entering the economy *endowed* with their human capital. On the other hand, (5.2.17) implies by usage of (5.2.12) also that

$$Y^{GDP}(t) = W_m(t) + W_o(t), \tag{5.2.18}$$

i.e. gross domestic product does represent the middle-aged and old's wealth *only*. While the existing cohorts' wealth is determined within the economy, the young's wealth is additively added to this – being created in the society's cradles. To illustrate this difference further, it is useful to interpret it in an intertemporal sense.

Intertemporal Interpretation

The intertemporal evolvement of $W(t)$ is driven by two separate concepts. Firstly, $W_m(t)$ as well $W_o(t)$ are defined via the returns on the corresponding previous savings of these cohorts, i.e. $S_y(t-1)$ and $S_m(t-1)$. The returns these savings earn are completely determined by the portfolio allocations and the equilibrium returns on both securities. Thus, these components comply perfectly with the standard intertemporal stock-flow concept in financial portfolio problems as in Samuelson [1969], Merton [1969*b*] or Fama [1970*b*]: the optimally reinvested part of the stock of wealth earns a return which determines its appreciated size in the subsequent period. In formal language this reads

$$\begin{aligned} \tilde{W}_m(t+1) &= (W_y(t) - C_y(t))\, \tilde{R}^*_{P,y}(t+1) \\ \text{and} \quad \tilde{W}_o(t+1) &= (W_m(t) - C_m(t))\, \tilde{R}^*_{P,m}(t+1) \end{aligned} \tag{5.2.19}$$

based on (3.4.7), (3.4.21) and (3.4.30). Observing the rate of changes in wealth implies tracking variations in a *stock* variable. The causes of this – the flow component from the production of output and distributed to human labor and physical capital in the form of rentals – is only reflected indirectly in the returns. Looking at the return definitions (3.2.13) and (3.3.9) verifies this. However, secondly, $W_y(t)$ does not follow this standard formalization of finance. The young's aggregated wealth is additively added to the intertemporally derived stocks of $W_m(t)$ and $W_o(t)$. From an aggregate perspective, this reflects the standard procedure to include endowment in addition to the reinvested wealth.[10] Yet, the value of this endowment is not chosen arbitrarily but determined by the anticipated labor income of the subsequent period. This discounting of future income flows is only possible in the financial concept of wealth, not in the economic accounting framework of gross domestic product or output. As $Y^{GDP}(t)$ represents all value that is created in a certain period t, it comprises only components that are *realized* in that period. The young's productive influence on the economy, however, becomes only effective when they work in period $t+1$. Then their wages will form a part of income and gross domestic product of *that* period. Hence, it must not be accounted for in period t.[11]

Fisher Separation versus General Equilibrium

To illustrate this further take Fig. 5.1, which is an extended version of Fig. 2.2 from Chap. 2. The first extension to the standard model of Fisher [1930] is the usage of Y^{GDP}.[12] Due to the introduced capital adjustment technology not output but gross domestic product is the driver of intertemporal transformation. Yet, based on (3.3.4) and the constant investment-output ratio, the difference between these two concepts is a constant factor

$$\begin{aligned} Y^{GDP}(t) &= Y(t) + \frac{1-\vartheta}{\vartheta} I(t) = Y(t) + \frac{1-\vartheta}{\vartheta} \Psi Y(t) \\ &= \left(1 + \frac{1-\vartheta}{\vartheta} \Psi\right) Y(t). \end{aligned} \tag{5.2.20}$$

This explains the relationship of the Y^{GDP}- and Y-curves in Fig. 5.1. As in the standard model, agents could *theoretically* extend their joint period

[10] Compare this, for instance, with the inclusion of labor income as discussed in Sect. 2.2.2.

[11] Furthermore, the framework does not contradict the notion that a closed economy on aggregate cannot accumulate wealth either. The human capital securities merely allow an intertemporal openness and thus the financial anticipation of future output generation.

[12] Furthermore, note that due to the certainty of the size of next period's labor force, $L(t+1) = N_m(t+1) = N_y(t)$, the production possibility set can be characterized solely by the investment amount.

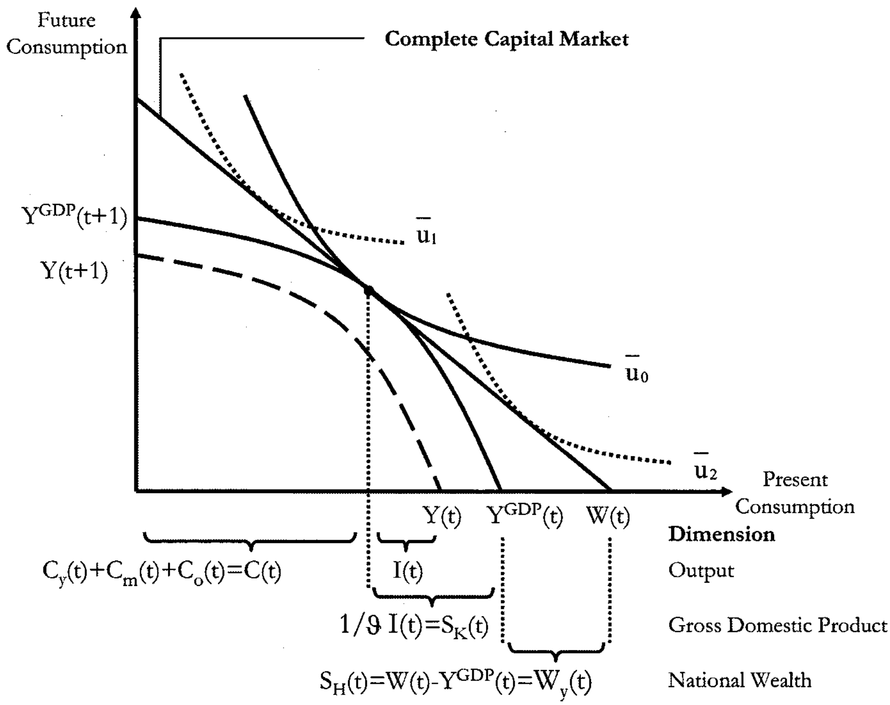

Fig. 5.1. Overlapping generations separation

Remarks: Tradability of human capital allows the saving generations to share investments in human and physical capital.

t-consumption, $C(t)$, beyond the maximal production of the production possibility set. In other words, they can realize any point on the line representing the *complete* capital market, like e.g. the tangential points to the indifference curves $\bar{u}_1$ or $\bar{u}_2$, according to the utility specification. At the extreme, the Fisher Separation allows the agents to consume all their discounted income, i.e. choose the point where the line representing the capital market intersects the horizontal axis. This means consuming the entire national wealth $W(t)$. However, in *equilibrium* consumption cannot exceed production. Hence, in aggregate the intertemporal decision must not contradict the feasibility restrictions from the production set. Without a real riskless security in positive net supply, riskless borrowing and lending among agents must sum to zero. Thus the tangential point of $\bar{u}_0$ is realized.[13] In this case $C(t)$ represents the

[13] $\bar{u}_0$ is merely to be interpreted as the indifference curve of the population representative agent and not as a social welfare function! Yet, with the simplicity of the framework – i.e. no differences within cohorts and consumption policies proportional to wealth – the distinction is not so severe.

general equilibrium amount of consumption independent from its split into $C_y(t)$, $C_m(t)$, and $C_o(t)$. $I(t) = Y(t) - C(t)$ is the amount of output dedicated to investments. In the dimension of gross domestic product this forgone consumption represents a value of $S_K(t) = 1/\vartheta\, I(t)$ due to the adjustment costs. Realizing the tangential point implies that period t's middle-aged work in t to generate $Y^{GDP}(t)$ and period t's young generation works in $t+1$ to produce $Y^{GDP}(t+1)$. In doing so they will receive the labor's share of output of that period.

However, the existence of the human capital security on the capital market allows the young to trade this future wage against consumption in period t and against participation in ownership of real capital securities. Since these two types of securities are in positive net supply the aggregated effect of them does *not* cancel out. The investors into the human capital securities sold by the young are the middle-aged agents. According to their portfolio policies, it is optimal to diversify their investments and allocate it partially in human capital securities. Basically this allows the young to consume already in period t - even though their endowment with human capital produces an economic flow only in period $t+1$. At the same time, it enables the middle-aged to deviate from the production optimal point of aggregated investments into real capital. In this sense, a kind of Fisher Separation of the aggregated savings and investment decision under risk reemerges. Due to the "*Overlapping Generations Separation*" the middle-aged are not limited to realize investments in physical capital only and the young are not forced to save their entire endowment in form of human capital securities. Instead, only the aggregated amount across all generations matters for the equilibrium.

Accounting Analogy

An analogy with standard business accounting might ease the interpretation. In the framework of the International Accounting Standards, for instance, all resources from which future economic benefits are expected to flow to a company are considered as its assets. Hence, the firm must account for them in its balance sheet and report changes of their value in the profit and loss statement.[14]

Considering the entire economy as such a firm, the stock of physical capital and the size of the labor force are its "assets" carried from one period to another. They are "acquired" on the output and human capital market and allow to generate the subsequent period's gross domestic product. The size of this product – the economy's revenues – is determined by the realizations of the production function's stochastics. As there are no economic profits in this neoclassical setting, the revenues are exactly offset by the costs of producing it; i.e. the rental payments to labor and physical capital add up exactly to the

[14]See IAS-Framework F. 49(a). Certainly the rules for capitalizing specific assets and fair value accounting are more complicated than presented here, but not due to the underlying economic principles.

size of the gross domestic product.[15] This distribution to the input factors determines how the claims on the firm's liability side change. On the opening balance sheet ownership of the total assets, i.e. all physical and human capital, is split between the agents of the middle-aged and the old according to each cohorts' allocation of savings in the previous period. This allocation has been realized by two types of shares, each participating in the rental of a single production factor. The distribution of the rentals allows to calculate the returns realized on each kind of shares. Hence, it determines the relative sizes of claims – or wealth – the two cohorts have on the the gross domestic product.[16]

Now a new generation enters the scene and joins the economy-corporation by a capital increase. On the asset side, the new young generation contributes the obligation to perform human labor in the subsequent period, which the economy-company accounts for as human capital stock. On the balance sheet's liability side, this adds the entry of the young's wealth. In order to comfort the consumption preferences of its owners, the economy realizes a kind of share repurchase program, partially distributing the generated consumable output and resulting in a contraction of the balance sheet: the old want to consume their entire stake and are thus entirely canceled from the liability side. Young and middle aged decide to consume their stake only partially. This share repurchase program reduces the size of the balance sheet from total wealth to total savings.[17]

Finally, young and middle-aged engage on the market for rental-shares and based on their diversification motives end up with an identical allocation. For the economy-corporation this trading among its owners has no effect: on the closing balance sheet's asset side there are still the stock of physical capital and the endowed labor force for next period. However, ownership of the pure-rental shares has changed among young and middle-aged agents.

Now the cycle restarts: the young become the middle-aged and the middle-aged retire to form the new old generation. Thus, in the new period's opening balance sheet ownership of the total assets is again split between these generations. Another accounting analogy shall be pointed out: the intertemporal

[15]This reflects the functional distribution of income and Euler's adding-up theorem illustrated in (2.2.15) of Sect. 2.2.5.

[16]Equation (5.2.18) shows this in case of doubt. Furthermore, note that $Y^{GDP}(t)$ indirectly reflects the anticipated value of future physical capital as seen in (3.3.4).

[17]This can by seen by manipulating (5.2.14) to get

$$W(t) - C(t) = S(t) = S_K(t) + S_H(t)$$

on the asset side. And from (3.5.2)

$$S(t) = S_y(t) + S_m(t)$$

on the liability side. Furthermore, notice that this is only an illustrative decomposition of the asset and liability side. Actually, everything takes place simultaneously.

determination of wealth corresponds to fair value accounting because it is based on realized market values. Accounting for the the young's contribution at the recapitalization requires to discount their risky future labor income. Yet, in the perfect world of this framework, the economy's accountants get the discount rate exactly right and obtain the fair value of this intangible asset.

5.2.4 Replication with PAYGO

With this understanding of the actual relations in the underlying unconstrained economy, one can now turn to the replication in the mimicking economy. Since there are no human capital securities in it, the analogy is not trivial.

National Income and Wealth

As mentioned earlier the lack of human capital securities limits savings to the physical capital kind, so that the identity of savings and investments is distorted only by the adjustment costs:

$$S'(t) = S'_K(t) = \frac{1}{\vartheta'} I'(t). \tag{5.2.21}$$

Furthermore, there is *no* need to introduce some government expenditures $G'(t)$ in the national income identity as possibly expected for the case with the public sector. Due to $i'_y(t)I'(t) + i'_m(t)I'(t) = I'(t)$ from (4.3.25) and the definition of $C'(t)$ in (4.3.16), all demand for consumable output units channeled through the government is already accounted for. To check this, split the components of the national income identity, (4.2.5), into

$$\begin{aligned} Y'(t) &= C'(t) + I'(t) \\ &= C'_y(t) + C'_m(t) + C'_o(t) + i'_y(t)I'(t) + i'_m(t)I'(t). \end{aligned} \tag{5.2.22}$$

One could now substitute some $G'(t)$ for $C'_y(t) + i'_y(t)I'(t)$,

$$Y'(t) = C'_m(t) + C'_o(t) + i'_m(t)I'(t) + G'(t),$$

but this does not yield much intuition apart from making it superficially more consistent with standard income accounting. The reason for the redundancy of $G'(t)$ is that the public sector does not consume on its own by expenditure in public infrastructure, education, defense etc. Instead, it has only a redistributive role between the generations.[18]

[18] Of course one would also like to include $C'_o(t)$ in $G(t)$. Yet, consumption of the old is only partially financed by the PAYGO system as they have private funded savings as well, so that this inclusion would be incorrect. It is not shown here because (5.2.22) is the better representation anyway.

This redistribution also manifests itself in the concepts of national wealth and gross domestic product. Since (4.2.27) established wealth with *gross* consumption, the equivalent to it on an aggregate level is

$$W'(t) = (1 + \tau_C)C'(t) + S'(t), \tag{5.2.23}$$

which is derived by applying the definitions of Sect. 4.3.2. The redistributive character of the consumption tax is reflected by the fact that it can be canceled out: since $W'_y(t) = N'_y(t)w'_y(t)$ due to (4.3.15), $w'_y(t) = \mathfrak{t}_y(t)$ as established in (4.2.28), and $\mathfrak{t}_y(t)N'_y(t) = \tau_C \times C'(t)$ by (4.2.22) one has

$$W'_y(t) = \tau_C \times C'(t). \tag{5.2.24}$$

Hence, one can subtract it on both sides of (5.2.23) and is left with

$$W'_m(t) + W'_o(t) = C'(t) + S'(t) = Y'^{GDP}(t), \tag{5.2.25}$$

where the equality with gross domestic product follows from $S'(t) = S'_K(t)$ and (4.3.22). This is the same result as in (5.2.18) for the unconstrained case: gross domestic product reflects only the wealth of the middle-aged and old. These two generations share ***primarily*** the value generated in a given period. The middle-aged work, receive thus labor's share of income and transfer it partially to the old via PAYGO. Both generations participate directly in physical capital's share of income based on their previous savings. Hence, – in a first step – gross domestic product is entirely distributed among these two cohorts. However, on the expenditure side the taxation of consumption redistributes – in a second step – a part of gross domestic product to the young generation so that they, too, can consume and invest in real capital. Combining (5.2.23) and (5.2.25), one can write

$$W'(t) = C'(t) + S'(t) + \tau_C C'(t) = Y'^{GDP} + \mathfrak{T}_y(t), \tag{5.2.26}$$

which is similar to (5.2.14) and relates national wealth and gross domestic product. Nota bene: the total transfers to the young take the role of savings allocated to human capital, which are zero in the constraint case. While the willingness of all generations to invest in human capital determined the size of the young's wealth before, it is now their propensity to consume that gives the young an initial wealth as seen in (5.2.24). This reallocation towards the young is one part of the replication procedure – necessary because of the nonexistence of the human capital securities.

The other part is the pay-as-you-go system which allows for the redistribution from the working middle-aged to the old cohort. To see its purely reallocating character, it is useful to analyze the primary distribution of Y'^{GDP} further. Starting with (5.2.25) and replacing the aggregate wealth variables by (4.2.29) and (4.2.30) – each multiplied with the respective cohort size – reads

$$
\begin{aligned}
Y'^{GDP}(t) &= W'_m(t) + W'_o(t) = N'_m(t)w'_m(t) + N'_o(t)w'_o(t) \\
&= N'_m(t)\left[s'_y(t-1)R'^{*}_{P,y}(t) + (1-\tau_\omega)\omega'(t)\right] \\
&\quad + N'_o(t)\left[s'_m(t-1)R'^{*}_{P,m}(t) + \mathfrak{t}_o(t)\right] \\
&= \left[S'_y(t-1)R'^{*}_{P,y}(t) + (1-\tau_\omega)L'(t)\omega'(t)\right] \\
&\quad + \left[S'_m(t-1)R'^{*}_{P,m}(t) + \mathfrak{T}_o(t)\right],
\end{aligned}
\tag{5.2.27}
$$

where the last lines follow from (4.2.11) and (4.2.19). Based on the assumption of a balanced PAYGO budget, i.e. (4.2.18) and (4.2.19), it holds that $\mathfrak{T}_o(t) - \tau_\omega L'(t)\omega'(t) = 0$. Using furthermore (4.3.17) for the portfolio returns, it remains that

$$
Y'^{GDP}(t) = \left(S'_y(t-1) + S'_m(t-1)\right) R'_K(t) + L'(t)\omega'(t). \tag{5.2.28}
$$

This can now be reformulated further using (4.3.20) and (4.3.21). With $S'_y(t-1) + S'_m(t-1) = S'_K(t-1) = K'(t)p'_K(t-1)$ and the results of (4.2.12) and (4.2.15) for $R'_K(t)$ and $L'(t)\omega'(t)$, it follows

$$
\begin{aligned}
Y'^{GDP}(t) &= K'(t)p'_K(t-1)\frac{\theta'(t)Y'(t) + \frac{1-\vartheta'}{\vartheta'}I'(t)}{K'(t)p'_K(t-1)} + (1-\theta'(t))\,Y'(t) \\
&= Y'(t) + \frac{1-\vartheta'}{\vartheta'}I'(t).
\end{aligned}
\tag{5.2.29}
$$

This is exactly the relationship that established gross domestic product in (4.2.9). In other words, the economic accounting framework is consistent and well defined.

The redistribution achieved by PAYGO can be observed by comparing the first line of (5.2.29) and the third line of (5.2.27): the primary distribution of $Y'^{GDP}(t)$ is to the macroeconomic input factors. Based on marginal productiveness real capital earns a rental of $\theta'(t)Y'(t) + (1-\vartheta')/\vartheta'\ I'(t)$ and human labor earns $(1-\theta'(t))Y'(t)$. While the old agents can participate in income of real capital due to their previous savings, $S'_m(t-1)$, they would have *no* claim on labor's share of income because they had not been able to invest in – nonexistent – human capital securities. Hence, the middle-aged would capture the entire payoff to human labor. But now the redistribution by the public pension scheme takes place. By paying the PAYGO contributions the working generation transfers a τ_ω-fraction of labor's share of output to the old. The calibration of the public sector's parameters in Chap. 4 achieves that this redistribution – and the one by the consumption tax – does not make any agent worse off than in the case with the complete capital market. Looking at the PAYGO redistribution as shown in (5.2.27) and (5.2.29) alone, could lead to the impression that the middle-aged generation is made worse off because it has to partially give up its claim on labor's share of output. However, its utility also includes the expected utils from consumption when this cohort is old. In this later period, they will be entitled to the PAYGO benefits in addition to their portion of physical capital's share of output and are thus made better of. The calibration to the results of the benchmark case implies that these two effects cancel out over the life cycle as defined by the utility rationale.[19]

Note furthermore that the proposed pension scheme does not accumulate any implicit debt. Each period it is fully funded and runs a balanced budget!

PAYGO Return

As indicated earlier, the partial redistribution of labor's share of income from the middle-aged to the old agents also transfers the associated risk to the retired generation. Because of the uncertainty of $\tilde{\theta}'(t)$ the wage rate and thus the total size of the contributions are risky. The old are confronted with this risk, since deficits are ruled out and the contribution rate is fixed and constant. While actual PAYGO schemes might have quite different objectives and exempt the old from this risk, this result is a direct consequence on the stringent focus on efficiency.[20] As a result of the riskiness of PAYGO benefits, the implicit return agents receive on the PAYGO participation is risky as well. Equation (4.2.21) established this return as

$$\tilde{\mathfrak{R}}(t+1) = \tilde{\omega}'(t+1)\omega'(t)\frac{N_y'(t)}{N_y'(t-1)}.$$

By combining the results of Chaps. 3 and 4 this PAYGO return can be linked directly to the underlying return on human capital, $\tilde{R}_H(t+1)$: inserting (4.2.13) for the wage rates and noting that the demography assumptions imply $N_m'(t) = N_y'(t-1)$ as well as $N_m'(t+1) = N_y'(t)$ allows to restate the finding for $\tilde{\mathfrak{R}}(t+1)$ as

$$\tilde{\mathfrak{R}}(t+1) = \frac{\frac{(1-\tilde{\theta}'(t+1))\tilde{Y}'(t+1)}{N_m'(t+1)}}{\frac{(1-\tilde{\theta}'(t))Y'(t)}{N_m'(t)}}\frac{N_y'(t)}{N_y'(t-1)} = \frac{1-\tilde{\theta}'(t+1)}{1-\theta'(t)}\frac{\tilde{Y}'(t+1)}{Y'(t)}. \quad (5.2.30)$$

In the nominator the risky payoff to human labor determines the future size of the PAYGO budget, in the denominator labor's share of output in the present period determines total contributions. Since the contribution rate is fixed, it

[19]This is one of the major advantages of the replication methodology: the hypothetical introduction of the human capital securities allows to establish and *solve* the trade-off of consumption over the life-cycle because the non-linearity problem is circumvented.

[20]Section 5.4 will focus on this difference.

cancels out and does thus *not* appear in (5.2.30). Now turn to the complete markets' case and note that one can solve the equilibrium return on human capital securities, (3.5.24), namely

$$\tilde{R}_H(t+1) = (1-\tilde{\theta}(t+1))\frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{1-\bar{\theta}}\frac{\tilde{Y}(t+1)}{I(t)},$$

for labor's future share of output:

$$(1-\tilde{\theta}(t+1))\tilde{Y}(t+1) = \frac{(1-\bar{\theta})I(t)}{(1-\vartheta)\Psi + \vartheta\bar{\theta}}\tilde{R}_H(t+1). \qquad (5.2.31)$$

Using this in (5.2.30), replacing $Y(t)$ by $I(t)/\Psi$, and canceling allows finally to link $\tilde{\mathfrak{R}}(t+1)$ tightly to $\tilde{R}_H(t+1)$:[21]

$$\begin{aligned}\tilde{\mathfrak{R}}(t+1) &= \frac{1}{1-\theta(t)}\frac{1}{Y(t)}\frac{(1-\bar{\theta})I(t)}{(1-\vartheta)\Psi + \vartheta\bar{\theta}}\tilde{R}_H(t+1)\\ &= \frac{(1-\bar{\theta})\Psi}{(1-\vartheta)\Psi + \vartheta\bar{\theta}}\frac{1}{1-\theta(t)} \times \tilde{R}_H(t+1). \qquad (5.2.32)\end{aligned}$$

The return on PAYGO and its risk characteristics are driven by the return on human capital because all uncertainty in (5.2.32) is due to the risky $\tilde{R}_H(t+1)$; the other elements on the right hand side are fixed or known at time t. Equation (5.2.32) is a strong finding: the financial performance of PAYGO cannot escape from economic reality. When the return on human capital in the underlying benchmark economy is high, so is the implicit return of the pay-as-you-go scheme. If the human capital securities would perform poorly so does the replicating PAYGO system.

Equation (5.2.32) would also allow the embedding of PAYGO returns into the standard portfolio problem. While Chaps. 3 had solved the portfolio rules with the hypothesized human capital securities, Chap. 4 introduced the public pension scheme. With the calibration derived there, (5.2.32) would allow to substitute it for the tradable human capital securities with adequate adjustments for the non-stochastic scaling factor. In this sense, this is the *indirect* derivation of the middle-aged agents' portfolio comprising returns from PAYGO and funded savings.

5.2.5 Growth Theory

Another important result of economic theory concerns the accumulation of physical capital and its implications for economic growth. Chapter 2 required the proposed answer to the pension problem to consistently integrate this aspect. This is necessary because the savings behavior of entire cohorts is not marginal.

[21] Due to the assumptions and conditions on replication the change to the notation without primes is easily possible and should not be confusing.

Physical Capital Accumulation

The aspect of physical capital accumulation is fully integrated in the benchmark as well as in the replication. Both economies are characterized by the same accumulation technology specified in (3.2.7) and (4.2.6) respectively:

$$K(t+1) = B\,I(t)^{\vartheta}\,K(t)^{1-\vartheta}.$$

While the analysis focuses again on the unconstrained case, it also holds for the mimicking economy. Based on (3.2.3) unconsumed output in the amount of $I(t)$ is dedicated to the accumulation of physical capital. Due to the adjustment costs the valuation of this as an input factor to the accumulation process is $S_K(t) = 1/\vartheta\ I(t)$.[22] As investment, unconsumed output increases the subsequent period's physical capital stock based on the above accumulation process.
The specification of (3.2.7) ensures that all the desired features of a capital accumulation process with convex adjustment cost are captures, i.e. a positive but decreasing marginal effect of investments:[23]

$$\frac{\partial K(t+1)}{\partial I(t)} = \vartheta B \left(\frac{K(t)}{I(t)}\right)^{1-\vartheta} > 0, \tag{5.2.33}$$

$$\frac{\partial^2 K(t+1)}{\partial I(t)^2} = -(1-\vartheta)\frac{\vartheta B}{I(t)}\left(\frac{K(t)}{I(t)}\right)^{1-\vartheta} < 0. \tag{5.2.34}$$

This decreasing marginal effect of investments on the stock of physical capital reflects the convexity of adjustment costs as discussed in Sect. 2.2.4: it becomes more costly to obtain one additional unit of capital in the subsequent period the higher current investment is.

Consumption Technology

Similarly, the consumption technology of (3.2.1) implies decreasing marginal returns on its input factors. This nests the desired feature discussed in Sect. 2.2.5: rentals paid to physical capital and to human labor cannot grow ad infinitum but are linked to the underlying accumulation processes. There is a negative feedback on returns on both factors as marginal productivity declines. To show this formally, note first that increases in the physical capital stock or the labor force for next period have a positive effect on subsequent output, but at a diminishing rate. Formally, insert (3.2.7) and (3.3.6) in (3.2.1) to obtain

$$\tilde{Y}(t+1) = \tilde{A}(t+1)B^{\tilde{\theta}(t+1)}I(t)^{\vartheta\tilde{\theta}(t+1)}K(t)^{(1-\vartheta)\tilde{\theta}(t+1)}N_y(t)^{1-\tilde{\theta}(t+1)}. \tag{5.2.35}$$

---

[22] This follows from (5.2.9).

[23] As all variables are positive and $0 < \vartheta < 1$ the derivatives' signs are obvious. The extreme cases $\vartheta = 0$ and $\vartheta = 1$ are excluded for this general analysis.

Taking the relevant derivatives yields[24]

$$\frac{\partial \tilde{Y}(t+1)}{\partial I(t)} = \vartheta\tilde{\theta}(t+1)\frac{\tilde{Y}(t+1)}{I(t)} > 0, \tag{5.2.36}$$

$$\frac{\partial^2 \tilde{Y}(t+1)}{\partial I(t)^2} = -\left(1-\vartheta\tilde{\theta}(t+1)\right)\vartheta\tilde{\theta}(t+1)\frac{\tilde{Y}(t+1)}{I(t)^2} < 0, \tag{5.2.37}$$

$$\frac{\partial \tilde{Y}(t+1)}{\partial N_y(t)} = \left(1-\tilde{\theta}(t+1)\right)\frac{\tilde{Y}(t+1)}{N_y(t)} > 0, \tag{5.2.38}$$

and
$$\frac{\partial^2 \tilde{Y}(t+1)}{\partial N_y(t)^2} = -\tilde{\theta}(t+1)\left(1-\tilde{\theta}(t+1)\right)\frac{\tilde{Y}(t+1)}{N_y(t)^2} < 0. \tag{5.2.39}$$

As expected for the Cobb-Douglas production function, it has positive but diminishing marginal products in either input factor.
One can now link returns on physical capital, human capital and – based on last section's findings – also indirectly on PAYGO to marginal increases in inputs. In equilibrium the returns on the securities are given by (3.5.23) and (3.5.24), i.e.

$$\tilde{R}_K(t+1) = \left[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)\right]\frac{\tilde{Y}(t+1)}{I(t)}$$

and
$$\tilde{R}_H(t+1) = \left(1-\tilde{\theta}(t+1)\right)\frac{(1-\vartheta)\Psi+\vartheta\bar{\theta}}{1-\bar{\theta}}\frac{\tilde{Y}(t+1)}{I(t)}.$$

The marginal effect of investments or increases in the young's cohort size onto the return on real capital is thus

$$\frac{\partial \tilde{R}_K(t+1)}{\partial I(t)} = \left[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)\right]\left(\vartheta\tilde{\theta}(t+1)-1\right)\frac{\tilde{Y}(t+1)}{I(t)^2} < 0, \tag{5.2.40}$$

$$\frac{\partial \tilde{R}_K(t+1)}{\partial N_y(t)} = \left[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)\right]\left(1-\tilde{\theta}(t+1)\right)\frac{\tilde{Y}(t+1)}{I(t)N_y(t)} > 0. \tag{5.2.41}$$

And for the human capital securities this yields

[24] Excluding again the extreme cases, all the inequalities are easily seen because of $0 < \tilde{\theta}(t+1) < 1$, $0 \leq \vartheta < 1$ and positive output, investments and cohort sizes.

$$\frac{\partial \tilde{R}_H(t+1)}{\partial N_y(t)} = \left(1-\tilde{\theta}(t+1)\right)^2 \frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{(1-\bar{\theta})I(t)} \frac{\tilde{Y}(t+1)}{N_y(t)} > 0, \tag{5.2.42}$$

$$\frac{\partial \tilde{R}_H(t+1)}{\partial I(t)} = \vartheta\tilde{\theta}(t+1)\left(1-\tilde{\theta}(t+1)\right) \frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{1-\bar{\theta}} \frac{\tilde{Y}(t+1)}{I(t)} > 0, \tag{5.2.43}$$

$$\frac{\partial^2 \tilde{R}_H(t+1)}{\partial N_y(t)^2} = -\,\tilde{\theta}(t+1)\left(1-\tilde{\theta}(t+1)\right)^2 \frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{(1-\bar{\theta})I(t)N_y(t)^2}\tilde{Y}(t+1) < 0. \tag{5.2.44}$$

Based on (5.2.32) it holds that $\partial\tilde{\mathfrak{R}}(t+1))/\partial\tilde{R}_H(t+1) > 0$, and therefore the findings for the return on human capital also hold for the return of the PAYGO pension scheme.

As expected, an increase in one input factor allows for higher returns in the other factor's security as seen in (5.2.41) and (5.2.43). But while human capital returns exhibit the expected positive but diminishing effect with respect to the amount of labor input, marginal increases in investments have a *negative* effect on return to physical capital!

There are two reasons for this difference. The first is that (5.2.40) and (5.2.42) measure two different things: while the former equation relates return of an input factor to *changes* in it, the latter relates returns to an input factor's level. For comparison one needs to establish the relationship of physical capital stock and returns on it. Using (3.5.23) and (5.2.35) one gets

$$\frac{\partial \tilde{R}_K(t+1)}{\partial K(t)} = \left[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)\right] \frac{(1-\vartheta)\tilde{\theta}(t+1)\tilde{Y}(t+1)}{I(t)K(t)} > 0. \tag{5.2.45}$$

Equation (5.2.45) corresponds to (5.2.42): ceteris paribus the higher the level of physical capital or human labor, the higher output will be in the following period and this allows for higher payoffs on both input factors. Yet, the decreasing marginal productivity in each input implies that these effects decline as the level of input factors increases. This can be seen from (5.2.40), (5.2.44) or equivalently from

$$\begin{aligned}\frac{\partial^2 \tilde{R}_K(t+1)}{\partial K(t)^2} &= \left[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)\right] \frac{(1-\vartheta)\tilde{\theta}(t+1)}{I(t)} \frac{\frac{\partial \tilde{Y}(t+1)}{\partial K(t)}K(t) - \tilde{Y}(t+1)}{K(t)^2} \\ &= \left[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)\right](1-\vartheta)\tilde{\theta}(t+1)\times \\ &\quad \left[(1-\vartheta)\tilde{\theta}(t+1) - 1\right]\frac{\tilde{Y}(t+1)}{I(t)K(t)^2} < 0.\end{aligned} \tag{5.2.46}$$

However, there is also a second effect causing $\partial\tilde{R}_K(t+1)/\partial I(t) < 0$, which can most clearly be seen in the case of full depreciation.

5.2.6 Full Depreciation

In the limiting case of fully depreciating real capital, the accumulation technology becomes trivial with $\vartheta = 1$ and $B = 1$ and (5.2.35) reduces to

$$\tilde{Y}(t+1) = \tilde{A}(t+1)I(t)^{\tilde{\theta}(t+1)}N_y(t)^{1-\tilde{\theta}(t+1)}. \tag{5.2.47}$$

While (5.2.45) and (5.2.46) would be zero because there is no accumulation of physical capital, the results of (5.2.36) to (5.2.44) remain unaffected. Hence, it still holds that higher investments imply lower returns to physical capital, but a larger young cohort increases the return on human capital. The cause for this are their different effects on the *intertemporal* terms-of-trade. This is visualized in Fig. 5.2.

A large size of the young cohort, which forms the next period's labor force, has the effect that the production possibility set increases at *every* level of invest-

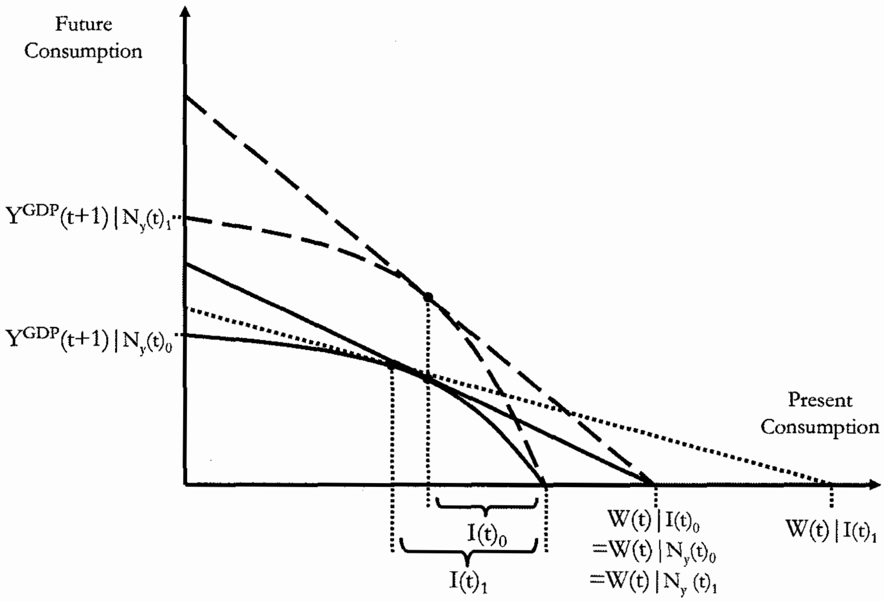

Fig. 5.2. Changes in input factors

Remarks: The figure illustrates the effects of changes in the macroeconomic input factors. A bigger young cohort, $N_y(t)_1 > N_y(t)_0$, extends future gross domestic product at *any* level of investments. On the other hand, an increase in investments, $I(t)_1 > I(t)_0$, changes the economy's equilibrium, increases the valuation of the young's human capital and thus translates into a higher aggregated wealth, $W(t)|_{I(t)_1} > W(t)|_{I(t)_0}$.

ments. Graphically, $Y^{GDP}|_{N_y(t)_0}$ shifts to $Y^{GDP}|_{N_y(t)_1}$ for $N_y(t)_1 > N_y(t)_0$.[25] However, aggregated wealth is the same for both sizes of the young generation, $W(t)|_{N_y(t)_0} = W(t)|_{N_y(t)_1}$! This is due to the fact that in the market for human capital securities there is *no* additional demand by the middle-aged because period t's cohort-aggregated allocations are unaffected by the increase in the young cohort's size. Hence, the market capitalization of human capital securities remains unchanged.[26] But since $N_y(t)_1 > N_y(t)_0$ implies a higher number of securities the same allocation is only achievable if the price per human capital security declines. This and the fact that human labor has a higher payoff in the next period – because output is larger and thus labor's share of it for any given realization of $\tilde{\theta}(t+1)$ – increases the return on human capital when the size of the young generation grows.

For a constant size of the young cohort but higher investments into physical capital, $I(t)_1 > I(t)_0$, the intertemporal effect is completely different: higher investments can only be realized by some parameter changes as investment and output are tied by Ψ. In other words, the move *on* $Y^{GDP}|_{N_y(t)_0}$ changes the nature of the realized equilibrium. Higher patience due to a larger subjective discount factor, for instance, can induce such higher savings. But with an unchanged intertemporal opportunity set, returns will be lower when investments increase. Though each additional unit yields more future output, it does so only at a decreasing rate. This effect, embodied in the Cobb-Douglas function of the consumption technology, has been formally described in (5.2.36) and (5.2.37). It means that the marginal return on physical capital declines, $\partial\tilde{R}_K(t+1)/\partial I(t) < 0$ as stated by (5.2.40). Furthermore, it can be shown that this decline dominates in the market return, so that $\partial\tilde{R}_M(t+1)/\partial I(t) < 0$.[27]

[25] Of course, in the special case $Y^{GDP}(t)$ equals $Y(t)$. Yet, since the effect also holds generally and is a second reason for (5.2.40), the general notation is used. It is given here in order to clearly separate it from the effect of physical capital stock, which is not present with full depreciation.

[26] In case of doubts refer to (3.5.31), which states

$$N_m(t+1)p_H(t)|_{\vartheta=1} = \frac{1-\bar{\theta}}{\bar{\theta}} I(t)\,.$$

The market value of human capital is proportional to the amount of investments, which remains unchanged for $Y^{GDP}|_{N_y(t)_0}$ or $Y^{GDP}|_{N_y(t)_1}$ in Fig. 5.2. The intuition is that physical capital per unit of labor is higher with a larger $I(t)$ making each unit of human capital more valuable. According to (3.5.17) this remains true for the case with accumulation.

[27] From $\tilde{R}_M(t+1)$ as in Table 3.4 one gets

$$\begin{aligned}\frac{\partial\tilde{R}_M(t+1)}{\partial I(t)} &= \left[(1-\vartheta)\Psi + \vartheta\bar{\theta}\right] \frac{\frac{I(t)\vartheta\bar{\theta}\tilde{Y}(t+1)}{I(t)} - \tilde{Y}(t+1)}{I(t)^2} \\ &= \left[(1-\vartheta)\Psi + \vartheta\bar{\theta}\right](\vartheta\bar{\theta}-1)\frac{\tilde{Y}(t+1)}{I(t)^2} \qquad < 0.\end{aligned}$$

Hence, future production is discounted at a lower rate resulting in a higher present value, which makes $W(t)|_{I(t)_1} > W(t)|_{I(t)_0}$ intuitive.
Another perspective on the different effects is that additional investments have opportunity costs in terms of forgone consumption, but young agents are an additional endowment to the economy. In Fig. 5.2 this is reflected by the fact that the production possibility set $Y^{GDP}|_{N_y(t)_1}$ clearly dominates $Y^{GDP}|_{N_y(t)_0}$: each $Y^{GDP}(t+1)$ can be achieved at less current investment. With the exogenous increase of young agents, the returns on each security have got to rise in order to preserve equilibrium at the given level of investments. If returns would not change, agents would have less incentive to save – and thus to realize $I(t)_0$ – because with the extended production possibility set they could achieve higher consumption in both periods.
Higher investments caused by changes in the parameters, on the other hand, result in a different intertemporal rationale of the agents: the opportunity costs of consumption today change so that agents want to consume less and save more at their new optimum, $I(t)_1 > I(t)_0$. Consequently, there will be higher demand on the market for human capital securities. Ceteris paribus, these yield a higher payoff as future output increases. But the supply of the human capital securities is fixed because the young cohort's size remains unchanged. Thus each of these securities is valued at a higher price, giving the young a larger endowment of wealth. Consequently, national wealth is higher, i.e. $W(t)|_{I(t)_1} > W(t)|_{I(t)_0}$ in Fig. 5.2.

Further Consistency Aspects

As consistency of this framework in the general case automatically implies the same for the special case of fully depreciating real capital, there is no point in redoing most crosschecks presented so far.
However, there are a few aspects to be emphasized: as indicated several times, national income accounting becomes simple when $\vartheta = 1$ and $B = 1$. Specifically the difference between the amount of unconsumed output units and its valuation as savings disappears. Based on (3.2.12) one has

$$S_K(t)|_{\vartheta=1} = I(t)|_{\vartheta=1} = K(t+1)|_{\vartheta=1}\,.$$

This is consistent with the definition of gross domestic product collapsing to output, which has already been established in Sect. 3.3.5. Hence, the standard identity of savings and investments into real capital requires only the adjustment for the portfolio allocations. This can be shown by first reducing (3.6.3) to the special case,

$$S(t)|_{\vartheta=1} = \frac{1}{\bar{\theta}} I(t)|_{\vartheta=1}, \tag{5.2.48}$$

and then inserting this and (3.5.33) into (5.2.8):

$$S_K(t)|_{\vartheta=1} = \alpha_K^M(t)|_{\vartheta=1} \times S(t)|_{\vartheta=1} = \bar{\theta} \times \frac{1}{\bar{\theta}} I(t)|_{\vartheta=1} = I(t)|_{\vartheta=1}\,. \tag{5.2.49}$$

Equation (5.2.49) is the same as using $\vartheta = 1$ in (5.2.9).

5.2.7 Illustrative Example Economy

Finally, as all numerical examples laid out so far have been drawn from the same example economy – in its benchmark and mimicking version – consistency checks for the initial ad-hoc assumptions can be performed. While this section focuses on the case with tradable human capital, the next will address the replication by PAYGO under the incomplete capital market.

Goods and Capital Markets

Starting with the goods market, remember that according to the results of Sect. 3.5.7 young as well as middle-aged agents allocate $\alpha_K^M(0) = 0.33$ of their savings in physical capital and $\alpha_H^M(0) = 0.67$ in human capital securities. This implies that the market capitalization of these securities are

$$\begin{aligned} K(1) \times p_K(0) &= \$606.06 \times 0.33 = \$200.00 \\ \text{and} \quad N_m(1) \times p_H(0) &= \$606.06 \times 0.67 = \$406.06. \end{aligned}$$

In the case without real capital accumulation, the consumption measured price of physical capital is one according to (3.2.16). As the physical capital stock of period $t = 1$ equals the previous period's investments one has

$$I(0) = \$200.00,$$

which is consistent with the assumption on investments in Sect. 3.2.6. In other words, the financial market of real capital securities clears, since investors' demand equals the financing demand of companies. Furthermore, based on the certainty assumption for the demography, $N_m(1) = N_y(0)$, the price of a single human capital security is

$$p_H(0) = \frac{\$406.06}{100} \approx \$4.06\,.$$

This is consistent not only with the initial ad-hoc assumption on $p_H(0)$ in Sect. 3.2.6 but also – and more importantly – with the risk-adjusted present value representation of future labor income in Sect. 3.5.7.

Investment-Output Ratio

Another crosscheck can be conducted using the investment-output ratio $\Psi = 6.04\%$ found in Sect. 3.6.5. Applying this ratio in (3.5.31) one can calculate the aggregated wealth endowment of the young cohort in period 1 for every realization of output $Y(1)$ from Sect. 3.2.6. In addition to this, the return and portfolio calculations from Sect. 3.3.6 and Sect. 3.5.7 allow to determine the wealth of the middle-aged and old generation in each of these states in period $t = 1$. Based on the optimal policies of Sect. 3.4.6 one can then derive their

saving and consumption decisions for that period.
For the realization of $Y(1)_1 = \$1,148.70$, for instance, aggregated wealth of *that* period's young cohort is given according to (3.5.31) by[28]

$$W_y(1)_1 = N_y(1)p_H(1) = \frac{1-\bar{\theta}}{\bar{\theta}}\Psi Y(1)_1 = \frac{1-0.33}{0.33} \times 6.04\% \times \$1,148.70$$
$$\approx \$140.97.$$

In the future state of $Y(1)_1$ the middle-aged and old would have a budget of

$$\begin{aligned} W_m(1)_1 &= s_y(0)N_y(0)\left[\alpha^*_{K,y}(0)R_K(1)_1 + \alpha^*_{H,y}(0)R_H(1)_1\right] \\ &= \$1.99 \times 100 \times [0.33 \times 115\% + 0.67 \times 226\%] \qquad \approx \$376.96 \\ \text{and} \quad W_o(1)_1 &= s_m(0)N_m(0)\left[\alpha^*_{K,m}(0)R_K(1)_1 + \alpha^*_{H,m}(0)R_H(1)_1\right] \\ &= \$3.26 \times 125 \times [0.33 \times 115\% + 0.67 \times 226\%] \qquad \approx \$771.74. \end{aligned}$$

Applying each generation's consumption policy implies that savings in period t are

$$\begin{aligned} S_y(1)_1 &= W_y(1)_1 \times \frac{0.6^2 + 0.6}{0.6^2 + 0.6 + 1} \approx \$69.05 \\ \text{and} \quad S_m(1)_1 &= W_m(1)_1 \times \frac{0.6}{0.6 + 1} \qquad \approx \$141.36. \end{aligned}$$

Again, the new old consume only. With an unchanged $\alpha^M_K(1) = \mathbb{E}_t[\tilde{\theta}(2)] = 0.33$ total allocation to physical capital, i.e. investments in period $t = 1$, is hence given by

$$I(1)_1 = 0.33 \times \$69.05 + 0.33 \times \$141.36 \approx \$69.43.$$

Lastly, dividing this by the output corresponding to this state gives the established

$$\Psi = \frac{I(1)_1}{Y(1)_1} = \frac{\$69.43}{\$1,148.70} \approx 6.04\%.$$

Since the same consistency can be established for the other realizations of $Y(1)$, the investment-output ratio is indeed constant, regardless of the realization of the random parameters! This illustrates the important claim of Sect. 3.6.2 on a numerical basis.

[28]Note that the overlapping generations framework implies that the formerly young are in period $t = 1$ the middle-aged, the formerly middle-aged are the old and the formerly old have passed away. The young of period $t = 1$ do thus correspond to an unborn cohort of period $t = 0$.

5.2.8 Replication of Illustrative Example Economy

Having shown the framework's consistency for the numerical example in case of tradable human capital, one can extend this exercise to the replication with the incomplete capital market. While Sect. 4.4.6 has explained that and how the replication works for the initial period $t = 0$, this section will extend it for a future state. Given Chap. 4's claim that the replication works dynamically this is an important supplement in the illustration of the numerical example.

Consumption in Period $t = 1$

To do so one has to extend the last section's analysis to the explicit first-best allocations in period $t = 1$ first. For the case of simplicity the illustration will again be limited to the future state of $Y(1)_1$ only. Yet, it is easy to realize the same calculations for the other states and proof so the consistency of the replication and the dynamic independence of the solution.
As the aggregated wealth budgets of all generations in period $t = 1$ have just been established, the corresponding allocations can be derived straightforwardly. The young of period $t = 1$ have an aggregated wealth of $W_y(1)_1 = \$140.97$. Assuming now a cohort size of $N_y(1) = 80$,[29] wealth for each individual young agent is in this state

$$w_y(1)_1 = \frac{W_y(1)_1}{N_y(1)} = \frac{\$140.97}{80} \approx \$1.76.$$

Since a young agent follows the same consumption-savings policy as derived in Sect. 3.4.6, his consumption will be

$$c_y^*(1)_1 = \frac{w_y(1)_1}{\beta^2 + \beta + 1} = \frac{\$1.76}{1.96} \approx \$0.90.$$

Similarly, based on the previous finding for the cohort wealth and his policy, a middle-aged agent of period $t = 1$ consumes

$$c_m^*(1)_1 = \frac{1}{\beta + 1} \frac{W_m(1)_1}{N_m(1)} = \frac{1}{1.60} \frac{\$376.96}{100} \approx \$2.36.$$

Because the old do not save, one has

$$c_o^*(1)_1 = \frac{W_o(1)_1}{N_o(1)} = \frac{\$771.74}{125} \approx \$6.17$$

[29] The size of the young cohort in period $t = 1$ is an additional assumption which has not been introduced yet. Since the previous cohort sizes have been specified as $N_y(0) = 100$, $N_y(-1) = 125$, and $N_y(-2) = 156$ in Sect. 3.6.5, $N_y(1) = 80$ supplements the description of a declining population at a yearly rate of about 0.75%.

for an old agent's consumption in the state of $Y(1)_1$. It is important not to make the mistake of comparing \$2.36 and \$6.17 for an assessment of agents' consumption smoothing. The true life-cycle patterns manifest themselves only in the relations of a given cohort. Considering $c_y^*(0) = \$2.07$ and $c_m^*(1)_1 = \$2.36$ for the generation born in period 0 or $c_m^*(0) = \$5.43$ and $c_o^*(1)_1 = \$6.17$ for the one born in $t = -1$ makes the development of individual consumption rather smooth and intuitive. However, to assess the full dimension of consumption smoothing one would also have to address the states characterized by output levels $Y(1)_2$, $Y(1)_3$ and $Y(1)_4$ from Sect. 3.2.6. Furthermore note that the higher consumption of old than of middle-aged reflects the no-bequest assumption.

Replication of Allocations

Turning to the replication, the question is whether the agents can achieve above consumption levels with the PAYGO system as derived in Chap. 4. To show that this is indeed the case one must first establish the primary distribution of income in the mimicking economy for this future state. Section 4.3.5 has established that output in the first state is $Y'(1)_1 = \$1,148.70$ and thus indeed identical to $Y(1)_1$. The realizations of the total factor productivity and production elasticity of real capital are $A'(1)_1 = 10$ and $\theta'(1)_1 = 0.2$. Consequently, labor's and real capital's share of output are in this state

$$\begin{aligned} (1-\theta'(1)_1)Y'(1)_1 &= (1-0.2) \times \$1,148.70 = \$918.96 \\ \text{and} \qquad \theta'(1)_1 Y'(1)_1 &= \quad 0.2 \times \$1,148.70 = \$229.74. \end{aligned}$$

With a cohort size of $N_m'(1) = 100$ in the working cohort this translates into a wage rate of $\omega(1)_1 = \$918.96/100 \approx \$9.19$ and based on the contribution rate $\tau_\omega = 73.3\%$ from Sect. 4.4.6 into a PAYGO budget of $\mathfrak{T}_o(1)_1 = \tau_\omega(1 - \theta'(1)_1)Y'(1)_1 \approx \$672.87$. Since $i_y' = 26.8\%$ of investments have been made by the young , this cohort captures as middle-aged now $\$229.74 \times 26.8\% \approx \$61.52$ of physical capital's payoff. At the same time the old of period $t = 1$ receive $\$229.74 \times (1 - 26.8\%) \approx \$168.22$ as income from real capital.
Following (4.2.30), this and the PAYGO budget of \$672.87 translate – for each single agent from the $N_o(1) = 125$ old ones – into a budget of

$$w_o'(1)_1 = (\$168.22 + \$672.87)\frac{1}{125} \approx \$6.73.$$

Since all agents have to pay the consumption tax of $\tau_C = 9.0\%$ as derived in Sect. 4.4.6, an old agent can achieve a consumption of

$$c_o'^*(1)_1 = \frac{1}{1+9.0\%} \times \$6.73 \approx \$6.17$$

because he does not save. Based on the wage level, contribution rate and the middle-aged's share of physical capital, each agent in this cohort has a budget of

$$\begin{aligned} w'_m(1)_1 &= (1-\tau_\omega)\omega(1)_1 + \frac{i'_y\theta'(1)_1 Y'(1)_1}{N'_m(1)} \\ &= (1-73.3\%)\$9.19 + \frac{26.8\% \times \$229.74}{100} \approx \$3.08. \end{aligned}$$

With the investment-output ratio replicated at $\Psi' = 6.045\%$ the generation of middle-aged agents saves now $i'_m\Psi'Y'(1)_1 = 73.2\% \times 6.045\% \times \$1,148.70 \approx \$50.84$ for investment in real capital. Each individual member of this generation can hence consume

$$c'^*_m(1)_1 = \frac{1}{1+9.0\%} \times \left(\$3.08 - \frac{\$50.84}{100}\right) \approx \$2.36.$$

Wealth of a young agent is determined by the tax-based subsidy from the public sector. As will be verified shortly, let the total level of consumption be $C'(1)_1 = \$1,079.26$. This translates into a fiscal budget of $\mathfrak{T}_y(1)_1 = \tau_C C'(1)_1 \approx \$96.98$, so that each individual young agent receives a transfer of $\mathfrak{t}_y(1)_1 = \mathfrak{T}_y(1)/N'_y(1) \approx \$1.21$ in the $Y'(1)_1$-state of the economy. Because young agents are not exempt from the consumption tax and realize $i'_y = 26.8\%$ of real capital investments, each of them achieves a consumption of

$$\begin{aligned} c'^*_y(1)_1 &= \frac{1}{1+\tau_C} \times \left(w'_y(1)_1 - \frac{i'_y\Psi'Y'(1)_1}{N'_y(1)}\right) \\ &= \frac{1}{1+9.0\%} \times \left(\$1.21 - \frac{26.8\% \times 6.045\% \times \$1,148.70}{80}\right) \approx \$0.90. \end{aligned}$$

Comparing these results with the agents' consumption levels in the the case of tradable human capital as just established yields that all of them can realize *exactly* the same level with the proposed PAYGO system and consumption tax:

$$c^*_y(1)_1 = \$0.90 = c'^*_y(1)_1 \quad c^*_m(1)_1 = \$2.36 = c'^*_m(1)_1 \quad c^*_o(1)_1 = \$6.17 = c'^*_o(1)_1.$$

In order to complete the proof that the replication works also dynamically, it must only be shown that the assumed aggregate consumption level is indeed realized and that the goods market clears. This is an easy exercise since the above consumption levels per agent and the demographic structure in period $t = 1$ result in a total consumption of

$$\begin{aligned} C'(1)_1 &= N'_y(1) \times c'^*_y(1)_1 + N'_m(1) \times c'^*_m(1)_1 + N'_o(1) \times c'^*_o(1)_1 \\ &= 80 \times \$0.90 + 100 \times \$2.36 + 125 \times \$6.17 \approx \$1,079.26, \end{aligned}$$

which justifies above's figure. Furthermore, with aggregate investments in physical capital of $I'(1)_1 = \Psi'Y'(1)_1 = 6.045\% \times \$1,148.70 \approx \$69.43$ the market for consumption goods clears perfectly:

$$Y'(1)_1 = C'(1)_1 + I'(1)_1 = \$1,079.26 + \$69.43 \approx \$1,148.70.$$

One could now extend this exercise to the other states of period $t = 1$ and would find that the PAYGO and fiscal system as calibrated in Sect. 4.4.6 achieves replication of the unconstrained allocations in each individual state. This illustrates numerically that the second-best solution works not only for the initial period, but also dynamically – regardless of the realizations of the economy's stochastics!

Note, however the important difference to the concept of dynamic completeness in the sense of Black and Scholes [1973]: while their insight relies on instantaneous rebalancing of the duplicating portfolio, the replication here is based on *constant* parameters in the public sector. The reason for this is that the replication by PAYGO relies more on the idea of giving access to the original payoffs of human labor when they are not tradable than on a synthetically duplication of them as in option pricing. Since each realization of $\tilde{A}'(t)$ and $\tilde{\theta}'(t)$ translates into a different wage level with a complete as well as with an incomplete capital market, PAYGO achieves that agents have somehow access to this risky payoff. Here it was shown that this can be realized with two public sector systems and constant parameters.[30] However, if the capital market is less than completely imperfect, it is well possible that one can achieve the same result with the PAYGO system only – yet under a dynamic rule for the contribution rate. Considering (5.2.32) for this interpretation, it is noteworthy that the implicit return on PAYGO is a rather linear transformation of human capital's return.

Yet, before reflecting on this hedging perspective further, the subsequent section will discuss random technical changes as an important aspect of the underlying production relation and reflects on diversification in this direction.

5.3 Technological Progress

So far the analysis has shown that a pay-as-you-go pension scheme can substitute for tradability of human capital. In the corresponding framework technological advances have not been addressed explicitly. Yet, a plausible intuition is that PAYGO should also allow to diversify the risks of technological progress and demographic development in old-age provisions. To address the question whether PAYGO can indeed accomplish a diversification between these uncertainties, technological change must be better incorporated into the model. In order to focus on this idea this section develops briefly a *new*, additional model to quickly analyze the impact of technological progress. As it will be shown that there is *no* diversification potential between funded savings and PAYGO pension provisions concerning the uncertainties of technology and demography, there is *no* need for a fully developed framework. Instead a generalization of the work of Bourguignon [1974] and Merton [1975] on stochastic

[30] The fiscal system depends crucially on the economy-wide consumption level and since the propensity to consumption is different between generations it is driven by the allocation of payoffs between generations as well.

growth theory is sufficient to derive the implications of technological change for total factor incomes.

5.3.1 Stochastic Growth Theory

While the framework so far has heavily relied on the overlapping generations framework and hence on discrete time, the easiest way for showing that PAYGO cannot help with technological change is to employ the mechanics of continuous time. Therefore this section reduces the analytic framework to its core elements of growth theory and translates it into the continuous representation. The corresponding notational change can be realized without loosing the intuition built so far.[31]

General Framework

To incorporate technological progress restart at the core of neoclassical growth theory, the simple model of Solow [1956] and Swan [1956] as presented in Sect. 2.2.5. A natural continuous time adoption of it is

$$Y_t = F(A_t, K_t, L_t), \tag{5.3.1}$$

$$I_t = Y_t - C_t, \tag{5.3.2}$$

$$\text{and} \quad dK_t = (I_t - \delta K_t)dt. \tag{5.3.3}$$

Y_t is the instantaneous output of consumable units corresponding to $Y(t)$ in the discrete setting so far. Analogously, the stock of physical capital shall be denoted K_t, the input of human labor L_t, total consumption C_t and gross investments I_t, which is just an intuitive adoption of their discrete time equivalents $K(t)$, $L(t)$, $C(t)$ and $I(t)$ from the previous chapters. Since the detailed description of agents' behavior in the overlapping generations structure as well as the explicit modeling of financial markets are not necessary to investigate the impact of technological progress, further distinction concerning cohorts or generations is not necessary.[32] Instead, the economic framework can be reduced to the absolutely essential elements, making the adjustment cost specification unnecessary as well. Therefore, the instantaneous change in the physical capital stock shall be modeled in a straightforward way, with a

[31] Admittedly, it seems inconvenient to sketch yet another model in order to incorporate technological progress. However, this option has the great advantage that the second framework can be strictly focussed on the issue of technological progress and can thus be tailored in continuous time. As time-subscripts are only used for this brief continuous time model there should be no confusion with the notation.

[32] This is for good reason: the beauty of the overlapping generations framework is its simplicity allowing to aggregate many individual agents to very few cohorts. A continuous time adoption would eliminate this advantage by distinguishing cohorts only by dt. This would generate an infinite number of overlapping generations, which is analytically intractable.

linear relationship between gross investment I_t, depreciation at rate δ and net investment dK_t/dt. Hence, the process of physical capital accumulation, (5.3.3), can be considered as an application of (3.2.8).
Equation (5.3.1) is the production relation with $F(A, K, L, t)$ as a general neoclassical macroeconomic production function, homogeneous of degree one and satisfying the conditions described in Sect. 2.2.5. While in (3.2.1) and (4.2.4) $\tilde{A}(t)$ and $\tilde{A}'(t)$ have been specified as the random total factor productiveness in the spirit of Solow [1956, Eq. (13)], and thus capturing the effect of technological change, the technological level A_t is now an *explicit* input in the production function. As in the discrete time framework, output can either be consumed or saved; yet, savings equal now investments in physical capital. This is because the notion of human capital as capitalized future labor earnings is not applicable without the overlapping generations framework. Thus (5.3.2) reflects the national income identity, analogously to (3.2.3) or (4.2.5).

Exogenous Technological Change

Equation (5.3.1) is the most general form to explicitly incorporate technological progress, which enters as a third input factor. This additional input is the only way to explain the empirically observed long-run growth in output per person, while the basic model of Solow [1956] and Swan [1956] implies that per capita output ceases due to the dampening effect of diminishing returns. Since the consequences of technological changes are analytically equivalent to increases in the macroeconomic input factors, progress is typically classified into three categories: capital-saving, labor-augmenting, and unbiased progress. The first refers to inventions that allow to produce the same output with less physical capital and the associated production functions are called Solow-neutral based on Solow [1969]. When productivity of both input factors increases so that they can be saved proportionally, progress is considered unbiased and the underlying production relation is neutral in the sense of Hicks [1932]. Harrod [1939] defines innovations as neutral if the relative input shares remain unchanged for a given ratio of real capital to output. This labor-augmenting progress does not only match the empirical facts of an increasing wage rate and a rather constant return on physical capital best, but is also the only type consistent with a stable steady-state ratio of per capita physical capital. It plays thus a dominant role in growth theory. Assume therefore that technological progress is of this *Harrod*-neutral form.[33] Robinson [1938] and Uzawa [1961] have shown that Harrod-neutrality implies a production function of the form $Y_t = F(A_t \times L_t, K_t)$, where the labor-augmenting influence of A_t is evident.

[33] See Harrod [1948] and e.g. Barro and Sala-i Martin [2004, Sect. 1.5.3] for the necessity that technological progress must be labor-augmenting.

Per Capita Notation

As in the basic representation in Chap. 2, it is convenient to reexpress the model in per capita notation, yet adjusted to address the labor-augmenting effect of technological change. Therefore, consider $A_t \times L_t$ as the *effective* amount of labor – the quantity of human labor scaled by the technological level which defines its efficiency. $k_t \equiv K_t/(A_t L_t)$ is now the ratio of physical capital to effective labor and $f(k_t) \equiv F(1, k_t)$ the intensive production function – a strictly concave monotonic increasing function on the nonnegative real line.[34] Based on the simply policies of Chap. 3 let savings be a deterministic fraction s of output as in Sect. 2.2.5. With investments in physical capital as the only form for allocating these savings, the accumulation of real capital is hence characterized by

$$dK_t = s\, Y_t\, dt - \delta\, K_t\, dt. \tag{5.3.4}$$

The other two determinants of the production function are supposed to follow diffusion processes:

$$dA_t = \alpha\, A_t\, dt + \sigma_A A_t\, dz_A \tag{5.3.5}$$

$$dL_t = n\, L_t\, dt + \sigma_L L_t\, dz_L, \tag{5.3.6}$$

where α and n are the drift and σ_A, σ_L the volatility parameters. Let dz_A and dz_L be standard Wiener processes. This specification as geometric Brownian motions with drift rules out a negative technological level as well as a negative labor level for given initial values $A_0 > 0$ and $L_0 > 0$. Furthermore, note that due to the absence of the overlapping generations framework, uncertainty concerning the birth rate as reflected by $\tilde{N}_y(t)$ in Sect. 3.1.1 now translates directly into a stochastic size of the labor force. Equation (5.3.6) is able to incorporate the desired feature of uncertain demographic development with a decreasing population, since n can be negative.[35]

In order to apply the usual trick of reducing the dynamics of the multi-dimensional macroeconomic model (5.3.4) to (5.3.6) to an one-dimensional process of k_t, Itô's lemma must be used. As shown in Appendix D.2.1 changes in real capital per effective labor are then given by

---

[34]The assumed homogeneity of the production function allows to rewrite (5.3.1) in per capita form as

$$Y_t = \frac{A_t L_t}{A_t L_t} F(A_t L_t, K_t) = A_t L_t\, F(1, K_t/(A_t L_t)) = A_t L_t\, F(1, k_t) \;\; = A_t L_t f(k_t).$$

[35]Bourguignon [1974], Malliaris and Brock [1982, Ch. 3] or Merton [1975] use all similar specifications for the labor process. The specification of equation (5.3.5) is somehow ad-hoc; however, this is an obvious extension of the typically assumed deterministic constant growth models for A_t where σ_A would be zero.

$$dk_t = \left[sf(k_t) - k_t(\alpha + n + \delta) + k_t(\sigma^2_{A+L} - \sigma_{AL})\right] dt - k_t \left[\sigma_A dz_A + \sigma_L dz_L\right], \tag{5.3.7}$$

where $\sigma_{AL} \equiv \sigma_A \sigma_L \rho_{z_L, z_A}$ is the instantaneous covariance of dz_A and dz_L and $\sigma^2_{A+L} \boxminus \sigma^2_A + \sigma^2_L + 2\sigma_{AL}$ the variance of $dz_A + dz_L$.
Equation (5.3.7) characterizes the dynamics of the ratio of physical capital to effective labor with *two* sources of uncertainty: one in the labor and one in the technology process. Given the substitutive character that these two inputs have in a Harrod-neutral formulation of technological progress, the additive influence of their variances is not surprising. However, it is noteworthy that in the drift term of dk_t not the entire summed uncertainty of both processes, σ^2_{A+L}, enters. Instead, it is reduced by the covariance σ_{AL}. This is a kind of diversification effect with respect to the macroeconomic inputs: surprises concerning technology or labor have a less severe effect when they are negatively correlated. Consequently, when $\sigma_{AL} < 0$ the instantaneous expected accumulation of physical capital is augmented by $-k_t \sigma_{AL} > 0$, which implies a higher output level according to the properties of $f(k_t)$. In this sense human labor and technological progress do not only supplement each other in absolute terms but also in a risk consideration.
Obviously, the general specification of (5.3.7) nests well established results as special cases. This can be shown by taking various limits:

$$\lim_{\sigma_A \to 0} dk_t = \left[sf(k_t) - k_t(\alpha + n + \delta - \sigma^2_L)\right] dt - k_t \sigma_L dz_L, \tag{5.3.8}$$

$$\lim_{\substack{\alpha \to 0 \\ \sigma_A \to 0}} dk_t = \left[sf(k_t) - k_t(n + \delta - \sigma^2_L)\right] dt - k_t \sigma_L dz_L, \tag{5.3.9}$$

and

$$\lim_{\substack{\alpha \to 0 \\ \sigma_A \to 0 \\ \sigma_L \to 0 \\ \delta \to 0}} dk_t = \left[sf(k_t) - k_t n\right] dt. \tag{5.3.10}$$

While (5.3.8) is the formulation for a deterministic technological progress, (5.3.9) corresponds to the results of Bourguignon [1974] or Merton [1975] who both abstract from technological progress. Equation (5.3.10) is the model of Solow [1956] and Swan [1956] with neither technological progress nor uncertainty. It is clearly the continuous counterpart of (2.2.14).
Like in the certainty version, one would be interested in the properties of the steady state characterized by (5.3.7). Still, in the stochastic case analyzed here, steady state implies not a single point but a unique distribution for k_t towards which its stochastic process tends. Appendix D.2.2 shows that the results of Bourguignon [1974], Merton [1975] or Malliaris and Brock [1982] can be extended for a model *with* technological progress as introduced by (5.3.5). Hence, asymptotical convergence of real capital per effective labor is given.[36]

[36] Note that for the case of discrete time Mirman [1973] and Brock and Mirman [1972] have proved the existence of a stochastic steady state under very general conditions.

Competitive Factor Incomes

However, it is important to note the difference of the processes for k_t and K_t. While the first follows the stochastic relation described in equation (5.3.7), the latter is not stochastic at all! At each point in time K_t as well as A_t and L_t are known; consequently, macroeconomic output is completely determined by equation (5.3.1). With a saving rule only dependent on data available at time t and a constant depreciation rate, the evolvement of the physical capital stock is completely determined by equation (5.3.4). Hence, the sample path for K_t is differentiable, well defined and locally certain. Based on this insight, one can use the traditional neoclassic framework to derive the factor shares in competitive markets because they are also well defined.

Adjusting the explanations of Sect. 3.2.1 on the firms' profit maximizing behavior for the continuous time model requires to acknowledge that firms can now decide on *both* of their inputs, the level of physical capital and the level of human labor. This is due to the continuous time limit in which the decisions on investments – and thus on the physical capital stock – and on labor are effectively simultaneous. Hence, firms maximize

$$\max_{K_t, L_t} \Pi_t \equiv p_{Y,t} \times Y_t - \omega_t \times L_t - r_t \times K_t - \delta K_t, \tag{5.3.11}$$

where Π_t is the correspondent of $\Pi(t)$ and r_t the instantaneous rental which physical capital earns in the production technology. Without a separate real capital adjustment process (5.3.11) must also address depreciation δK_t. Again it is possible to normalize the price of consumption goods to one, i.e. $p_{Y,t} \equiv 1 \,\forall\, t$. First order conditions corresponding to this optimization imply that each input factor is employed up to the point where marginal productivity equals marginal costs:

$$\frac{\partial \Pi_t}{\partial K_t} = \frac{\partial F}{\partial K_t} p_{Y,t} - (r_t + \delta) = 0 \tag{5.3.12}$$

$$\text{and} \quad \frac{\partial \Pi_t}{\partial L_t} = \frac{\partial F}{\partial L_t} p_{Y,t} - \omega_t = 0 \tag{5.3.13}$$

with ω_t as wage rate analogous to $\omega(t)$ from the previous chapters.

Finally, note that the usage of k_t allows to restate the marginal products of physical capital, $\partial F/\partial K_t$, and human labor, $\partial F/\partial L_t$. As by definition of $f(k_t)$ it holds that $F(A_t L_t, K_t) = A_t L_t f(k_t)$, one can express the marginal products as

$$\frac{\partial F}{\partial K_t} = A_t L_t \frac{\partial f\left(\frac{K_t}{A_t L_t}\right)}{\partial K_t} = A_t L_t \frac{\partial f(k_t)}{\partial k_t} \frac{1}{A_t L_t} \\ = \frac{\partial f(k_t)}{\partial k_t} \tag{5.3.14}$$

and
$$\frac{\partial F}{\partial L_t} = A_t f(k) + A_t L_t \frac{\partial f\left(\frac{K_t}{A_t L_t}\right)}{\partial L_t} = A_t f(k) + A_t L_t \frac{\partial f(k_t)}{\partial k_t} \frac{-K_t}{(A_t L_t)^2} \\ = A_t \left[f(k_t) - k_t \frac{\partial f(k_t)}{\partial k_t} \right]. \tag{5.3.15}$$

Assuming again competitive factor markets and using the last formulations in (5.3.12) and (5.3.13) leads to well known results, here shown for the continuous time version:

$$r_t = \frac{\partial f(k_t)}{\partial k_t} - \delta \tag{5.3.16}$$

$$\text{and} \quad \omega_t = A_t \left[f(k_t) - k_t \frac{\partial f(k_t)}{\partial k_t} \right]. \tag{5.3.17}$$

The rental price of each factor reflect the corresponding marginal product. In the case of physical capital it is adjusted for the diluting effect of depreciation, and for human labor it is adjusted for the scaling effect of labor-augmenting progress. The standard neoclassical result reemerges as a special case when the technology level A_t is normalized to one and depreciation is ignored. Since Merton [1975] abstracts from technological change, this simplification is also present in his analysis.

5.3.2 Factor Incomes in Cobb-Douglas Case

Having established the required core of stochastic growth theory, one can now derive further results for various formulations of the macroeconomic production function $F(A_t K_t, L_t)$. In order to closely relate this analysis to the framework of Chaps. 3 and 4, assume again that it is of the Cobb-Douglas type. Thus let now

$$Y_t = (A_t L_t)^{1-\theta} K_t^{\theta}, \tag{5.3.18}$$

which closely mirrors (3.2.1) with $0 < \theta < 1$.[37] Yet, note the differences: while in the previous framework the production elasticity of real capital has been stochastic, uncertainty is now entering (5.3.18) *only* via the stochastic dynamics of A_t and L_t as specified by (5.3.5) and (5.3.6). Furthermore, there is no general random scaling parameter reflecting total factor productivity; instead

[37]The extreme cases of $\theta = 0$ and $\theta = 1$ are of no interest here because they eliminate the substitutive character of the two inputs.

A_t is incorporated as implied by Harrod-neutral technological progress achieving the goal of addressing this important aspect.[38] Furthermore, depreciation shall be ignored without loss of generality, i.e. $\delta = 0$.

Dynamics of k_t, r_t and ω_t

The above specification of the production function implies that its per capita formulation and its derivative are given by

$$f(k_t) = \frac{F(A_t, K_t, L_t)}{A_t L_t} = \frac{(A_t L_t)^{1-\theta} K_t^\theta}{A_t L_t} = \frac{K_t^\theta}{(A_t L_t)^\theta} = k_t^\theta \quad (5.3.19)$$

$$\text{and} \quad \frac{\partial f(k_t)}{\partial k_t} = \theta k_t^{\theta-1}. \quad (5.3.20)$$

Using these specifications in (5.3.7), the process of k_t in the Cobb-Douglas case without depreciation can be stated explicitly as

$$\begin{aligned} dk_t &= \left[sf(k_t) - k_t(\alpha + n + \delta) + k_t(\sigma_{A+L}^2 - \sigma_{AL})\right] dt - k_t \left[\sigma_A dz_A + \sigma_L dz_L\right] \\ &= k_t \left\{\left[sk_t^{\theta-1} - \alpha - n + \sigma_{A+L}^2 - \sigma_{AL})\right] dt - \left[\sigma_A dz_A + \sigma_L dz_L\right]\right\}. \quad (5.3.21) \end{aligned}$$

Based on (5.3.16) and (5.3.17), in the Cobb-Douglas case without depreciation wages and physical capital rentals are given as

$$r_t = \theta k_t^{\theta-1} \quad (5.3.22)$$

$$\text{and} \quad \omega_t = A_t \left(k_t^\theta - k_t \theta k_t^{\theta-1}\right) = (1-\theta) A_t k_t^\theta. \quad (5.3.23)$$

Applying Itô's lemma on these formulations – as done in Appendix D.2.3 – gives the dynamics for the rentals of physical capital and wages:

$$\begin{aligned} dr_t &= (\theta - 1) r_t \left\{ \left[sk_t^{\theta-1} - \alpha - n + \frac{\theta}{2}\sigma_{A+L}^2 - \sigma_{AL} \right] dt \right. \\ &\quad \left. - (\sigma_A dz_A + \sigma_L dz_L) \right\} \quad (5.3.24) \\ d\omega_t &= \omega_t \left\{ \alpha dt + \theta \left[sk_t^{\theta-1} - (\alpha + n) - \frac{1}{2}(1-\theta)\sigma_{A+L}^2 + \sigma_L^2 \right] dt \right. \\ &\quad \left. + \left[(1-\theta)\sigma_A dz_A - \theta \sigma_L dz_L\right] \right\}. \quad (5.3.25) \end{aligned}$$

In order to interpret these findings, note first that both the wage level and the real capital rental, are positive because all factors in (5.3.22) and (5.3.23) are

[38] However, note that the production function in Chaps. 3 and 4 does *not* contradict the proofs of Robinson [1938] and Uzawa [1961], even though specification there implies a Hicks-neutral interpretation of progress. The reason for this is that the Cobb-Douglas production function is both, Hicks-neutral *and* Harrod-neutral, due to its elasticity of substitution of one.

positive. Second, with $r_t > 0$, $\omega_t > 0$ and $0 < \theta < 1$, their expected changes exhibit some standard relationships:

$$\frac{\partial\,\mathbb{E}\,[dr_t]}{\partial s} = (\theta-1)k_t^{\theta-1}r_t dt < 0 \qquad \frac{\partial\,\mathbb{E}[d\omega_t]}{\partial s} = \theta k_t^{\theta-1}\omega_t dt > 0 \tag{5.3.26}$$

$$\frac{\partial\,\mathbb{E}\,[dr_t]}{\partial \alpha} = -(\theta-1)r_t dt > 0 \qquad \frac{\partial\,\mathbb{E}[d\omega_t]}{\partial \alpha} = (1-\theta)\omega_t dt > 0 \tag{5.3.27}$$

$$\frac{\partial\,\mathbb{E}\,[dr_t]}{\partial n} = -(\theta-1)r_t dt > 0 \qquad \frac{\partial\mathbb{E}[d\omega_t]}{\partial n} = -\theta\omega_t dt < 0 \tag{5.3.28}$$

The neoclassical paradigm of positive but diminishing marginal productiveness implies that a higher savings rate increases wages but decreases physical capital's rentals. Higher savings correspond to a higher capital stock, which – ceteris paribus – means that capital is abundant relative to human labor. Hence, it has a smaller expected rental increase while the wage level growth is higher as seen in (5.3.26). A higher drift α of technological progress allows a larger output for every level of real capital and human labor and has thus a positive effect on rental and wage growth. And based on (5.3.28), the higher the drift n in the population change the smaller are the expected increases in wages and the higher those of rentals – as expected due to the relative scarcity argument. In other words, the neoclassical standard features are embedded in this model of stochastic growth.

The influence of the uncertainties on the expected changes are of special interest:[39]

$$\frac{\partial\,\mathbb{E}\,[dr_t]}{\partial \sigma_A^2} = (\theta-1)\frac{\theta}{2}r_t dt < 0 \qquad \frac{\partial\mathbb{E}[d\omega_t]}{\partial \sigma_A^2} = (\theta-1)\frac{\theta}{2}\omega_t dt < 0 \tag{5.3.29}$$

$$\frac{\partial\,\mathbb{E}\,[dr_t]}{\partial \sigma_L^2} = (\theta-1)\frac{\theta}{2}r_t dt < 0 \qquad \frac{\partial\mathbb{E}[d\omega_t]}{\partial \sigma_L^2} = (\theta+1)\frac{\theta}{2}\omega_t dt > 0 \tag{5.3.30}$$

$$\frac{\partial\,\mathbb{E}\,[dr_t]}{\partial \rho_{z_A,z_L}} = (\theta-1)^2\sigma_A\sigma_L r_t dt > 0 \qquad \frac{\partial\mathbb{E}[d\omega_t]}{\partial \rho_{z_A,z_L}} = (\theta-1)\theta\sigma_A\sigma_L\omega_t dt < 0 \tag{5.3.31}$$

The inequalities of (5.3.29) show that the higher the variance in the technology process the smaller the expected changes in payoffs to each input factor. This is an intuitive result, since the production function (5.3.18) is concave in both input factors. Hence, a low realization in the technology process has a more severe negative effect on marginal productiveness at any level of real capital and human labor than a a high realization has positive one. Because each factor's compensation is tied to its marginal productiveness by (5.3.22) and (5.3.23), expected changes in rentals and wages follow this pattern. In this sense the concave macroeconomic production function exhibits uncertainty

[39] All inequalities are again implied by $r_t > 0$, $\omega_t > 0$ and $0 < \theta < 1$.

aversion analogously to the risk aversion of an investor with a concave utility function.

While uncertainty in the technology process has this parallel effects, population variance increases expected wage changes but decreases growth of real capital rentals as seen in (5.3.30). However, the reasoning is the same: a low realization of growth in the labor force in (5.3.6) makes physical capital relatively abundant and thus commanding a lower rental while human labor is scarce and thus highly paid. On the other hand, a high L_t generates a higher r_t and a lower ω_t. Yet, due to the concavity of the production function with respect to labor input the former effect dominates the latter, resulting in the relationships observed in (5.3.30).

Equation (5.3.31) implies that the higher the correlation of population and technology processes the higher the expected change in physical capital rentals is and the lower is the change in wages. Given the Harrod-neutral specification of technological progress this is an equally intuitive finding: if there is a high positive correlation between the processes of (5.3.5) and (5.3.6), high realizations in both process are likely to coincide, resulting in large availability of effective labor, A_tL_t. In this case physical capital is relatively scarce and earns thus a high rental, while human labor receives only a low wage rate. On the other hand, with a negative ρ_{z_A,z_L} low realizations in the labor process are associated with high realizations in technological progress, making the few units of actual human labor highly productive. Hence, in this case expected increases in wages would be higher and the rentals' growth lower.

Investigating finally the variance components of (5.3.24) and (5.3.25) points at some potential for diversification between population and technology uncertainty, since $(1-\theta)\,(\sigma_A dz_A + \sigma_L dz_L)$ and $[(1-\theta)\sigma_A dz_A - \theta\sigma_L dz_L]$ are far from perfectly correlated. Yet, in order to thoroughly address this question the *total* income paid to each input factor must be considered.

Total Factor Incomes

Having established the stochastic dynamics of wages and physical capital rentals, one can derive the corresponding processes easily. Let total factor income paid to real capital be denoted $Y_{r,t}$ and total income paid to human labor $Y_{\omega,t}$. As in Chaps. 3 and 4, each input's share of output is given by the product of input amount and input price. In other words, physical capital receives a total payoff of $Y_{r,t} = r_tK_t = r_tk_tA_tL_t$ and human labor $Y_{\omega,t} = \omega_tL_t$. Applying Itô's lemma on these tautologies yields the stochastic dynamics of each factor's total income:[40]

[40] Appendix D.2.3 shows the derivations.

$$dY_{r,t} = Y_{r,t}\left[\theta s k_t^{\theta-1} + (1-\theta)\left(\alpha + n - \frac{\theta}{2}\sigma_{A+L}^2 + \sigma_{AL}\right)\right] dt$$
$$+(1-\theta)Y_{r,t}(\sigma_A dz_A + \sigma_L dz_L) \tag{5.3.32}$$
$$dY_{\omega,t} = Y_{\omega,t}\left[\theta s k_t^{\theta-1} + (1-\theta)\left(\alpha + n - \frac{\theta}{2}\sigma_{A+L}^2 + \sigma_{AL}\right)\right] dt$$
$$+(1-\theta)Y_{\omega,t}(\sigma_A dz_A + \sigma_L dz_L). \tag{5.3.33}$$

This is a stunning result and finally proofs the claim that there is *no* diversification potential between funded savings and PAYGO pension provisions concerning the uncertainties of population change and technological progress. To see this, note that according to (5.3.32) and (5.3.32) the total income paid to each input factor follows *exactly* the same stochastic process – though on different absolute levels – and has thus the same random component as the total income of the other factor! But with identical randomness these two factor incomes are perfectly correlated and do thus *not* offer any possibility to diversify and eliminate some risk.

5.3.3 Implications for PAYGO

To focus further on this finding, remember from Chap. 3, that the return on the two securities of the replication framework has been determined by the respective future payoffs. Section 3.3 has made this link explicit by pointing out the present value interpretation. In order to repeat the idea of replicating tradable human capital by a pay-as-you-go pension system, one would now try to establish this section's continuous time equivalents of Chap. 3's $\tilde{R}_H(t+1)$ and $\tilde{R}_K(t+1)$. Such a $\tilde{R}_{H,t}$ should reflect the payoff to human labor, i.e. $Y_{\omega,t}$, and $\tilde{R}_{K,t}$ the payoff to physical capital, i.e. $Y_{r,t}$. Yet, since (5.3.32) and (5.3.32) show that these future payoffs are *perfectly* correlated, there is no point in deriving the entire replication framework, establishing equilibrium characteristics and then trying to calibrate the PAYGO allocations to the first-best solution. The fundamental result will not change: in the Cobb-Douglas setting, extended for Harrod-neutral technological progress as specified by (5.3.5), the payoffs of human and real capital are perfectly correlated with respect to the uncertainties of population change and progress. Therefore, the corresponding returns on securities tracking these factor incomes must be correlated equally perfect. And due to the identical drift components they even represent perfect substitutes for any risk-averse investor! But without a dedicated role for hypothetical human capital securities there is no replication role for PAYGO either! In other words, the naive notion that a public pension scheme allows retirees to participate in the stochastic dynamics of population while funded savings are driven by technological progress and thus a combination of these two offers some diversification, *cannot* be supported!

Given the Harrod-neutral formulation of technological progress this not a surprise. As neutrality of this form implies that technological advances are labor-augmenting, technological progress has the same consequences for macroeconomic production as changes in the population. The formulation of an intermediary input factor, the *effective* labor units A_tL_t, reflects this. In this meta-combination technology and population present close substitutes. Innovations in one of them can thus compensated or aggravated by the other. Of course, the introduction of technological process has different effects on the compensation of real capital and human labor on a per-unit consideration. This can clearly be seen when (5.3.25) is written as

$$d\omega_t = \theta\omega_t\left\{\left[sk_t^{\theta-1} - (\alpha+n) - \frac{1}{2}(1-\theta)\sigma_{A+L}^2 + \sigma_L^2\right]dt - (\sigma_A dz_A + \sigma_L dz_L)\right\}$$
$$+\omega_t\left(\alpha dt + \sigma_A dz_A\right). \tag{5.3.34}$$

Comparing (5.3.24) and (5.3.34) shows that the uncertainties $\sigma_A dz_A + \sigma_L dz_L$ enter both processes directly and corrected for the respective production elasticities. This corresponds to the expected effect caused by the decreasing marginal productiveness in each input factor. Yet, wages capture the technological progress in an additional separate term, $\alpha dt + \sigma_A dz_A$, reflecting the substitutive relationship of A_t and L_t for the effective labor units. Turning to the total factor incomes in (5.3.32) and (5.3.33) this finding changes: the different effects on the per-unit level are *exactly* offset by divergent evolvements of the total amount of each input factor. The net stochastics of the payoff per unit and the number of input units available are the same for physical capital and human labor. Of course, the absolute level of $Y_{r,t}$ and $Y_{\omega,t}$ will be different, but since they follow the same dynamics they offer the same risk-return pattern to a potential investor. As Chap. 2 emphasized that for a sustainable pension strategy an aggregated macroeconomic perspective must be considered, it is the processes of (5.3.32) and (5.3.33) for the total factor incomes that matter and not the per-unit ones (5.3.24) and (5.3.25).

Economic Profits

Having established the dynamics of each factor's total compensation one can easily sum them to the dynamics of national income. Since $Y_t = Y_{r,t} + Y_{\omega,t}$ one has

$$\begin{aligned} dY_t &= dY_{\omega,t} + dY_{r,t} \\ &= Y_t\left[\theta sk_t^{\theta-1} + (1-\theta)\left(\alpha + n - \frac{\theta}{2}\sigma_{A+L}^2 + \sigma_{AL}\right)\right]dt \\ &\quad +(1-\theta)Y_t(\sigma_A dz_A + \sigma_L dz_L). \end{aligned} \tag{5.3.35}$$

This equation shows again that all stochastics governing the macroeconomic output are equally reflected in labor's and real capital's share of output. Equations (5.3.32), (5.3.33), *and* (5.3.35) have all the same variance components relative to their levels. In other words, the uncertainties of the population and technology process influence national income to the same extent as they influence individual factor incomes. Each of them fully mirrors the total risk. Diversifying it partially away is hence not possible by any combination of them.

Furthermore, it is important to understand that this framework is cohesive. The assumptions of profit maximizing firms as given by (5.3.11) and competitive factor markets imply that economic profits are zero. In the presented Cobb-Douglas case without depreciation one has by (5.3.22) and (5.3.23) that

$$\begin{aligned} \Pi_t &= Y_t - K_t r_t - L_t \omega_t = A_t L_t k_t^\theta - A_t L_t k_t \theta k_t^{\theta-1} - L_t (1-\theta) A_t k_t^\theta \\ &= A_t L_t \left[k_t^\theta - \theta k_t^\theta - (1-\theta) k_t^\theta \right] = 0. \end{aligned} \tag{5.3.36}$$

Moreover, one can even show that it also holds that $d\Pi_t = 0$. Even though the economy is characterized by two sources of uncertainty profits are immutably zero. But due to the neoclassical paradigm rentals paid to each factor are *economic* ones, so that they already include any risk-adjustments.

Conclusion

With the additional little model this section has not only addressed technological progress in its most accepted kind of Harrod-neutrality explicitly, but also shown that this does not lead to diversification potential concerning risks of technological progress and demographic changes. Thereby the mechanics of continuous time made the presentation convenient as the stochastic dynamics can be derived straightforwardly.

Furthermore, it must be emphasized that uncertainty matters for the underlying growth model. The dynamics of output, factor incomes and physical capital per effective labor all include considerable adjustments for the modeled uncertainties – not only in the variances but in the drift components! Merton [1975, Table 2] shows that the certainty estimates are systematically biased. Thus care must be taken in using certainty results of as proxies for the first moments when ignoring the stochastics. This also implies that the findings of simulations on PAYGO systems must be interpreted cautiously, since the typically ignore uncertainty and are thus subject to these biases.

The elimination of any diversification leeway is due to the fact the human labor captures all the benefits of advances in technology. Hence, human capital unites the demographic and technological uncertainties in its capitalization of wages. If physical capital would capture the merits of technological progress, the dynamics might change and hence the implications for diversification and replication with PAYGO. Yet, the empirical evidence is in favor of Harrod-neutral progress and thus rejecting the naive diversification idea.

5.4 Design of PAYGO Schemes

With the additional insight on technological progress one can now unite the lessons from the last section and from the replication framework of Chaps. 3 and 4 into an integrated understanding of the pay-as-you-go mechanism. This will lead to important results with respect to diversification and for the design of PAYGO pension schemes.

5.4.1 Risk Diversification

First, note that the two models have not been so different. While it was shorter to show the nonexistent diversification potential of Harrod-neutral progress in the continuous time context, the underlying economics are still quite identical. Both have been based on a macroeconomic growth model with a Cobb-Douglas production function. Factor incomes have been derived analogously under the neoclassical paradigm of marginal productiveness. Yet, since in the model with explicit technological progress labor's and physical capital's share of income followed the same stochastic dynamics, there would have been no point to develop all the subsequent steps of assuming a human capital security and replicating it with PAYGO. In contrast to this, the factor incomes in the original model offered diversification potential leading to the special role of PAYGO as surrogate for tradable human capital.
The reason for this difference is rather obvious: while Sect. 5.3 has been based on a model with a *constant* production elasticity of real capital, θ, in the detailed framework of Chaps. 3 and 4 this has been a randomly *fluctuating* variable $\tilde{\theta}(t)$. All that PAYGO achieves is that the risk caused by this uncertainty is borne by all age groups – as it is with tradability.

Diversifiable Risk

An adequately designed pay-as-you-go mechanism serves thus as a devise for an intergenerational sharing of the risk concerning the *relative* sizes of physical capital's and human labor's share of national income. Factor income risk becomes diversifiable. Since, in the neoclassical setting, the distribution of income is driven by the marginal productiveness of each factor, it requires uncertainty in the marginal factor productivity to produce a diversification potential. For the Cobb-Douglas case it is possible to characterize the respective productiveness with a single variable since it is assumed that the production elasticities add up to one as discussed in Sect. 3.2.1. Thus the random $\tilde{\theta}(t)$ – identified as real capital's production elasticity – is sufficient to produce the required uncertainty in the factor incomes. With its fixed θ the continuous time model, on the other hand, lacked this uncertainty and resulted thus in identical dynamics for $Y_{r,t}$ and $Y_{\omega,t}$.
In the replication model these factor incomes have been referred to as payoffs in order to facilitate the interpretation of the corresponding securities. As

summarized in Table 3.1 each of the $K(t+1)$ real capital securities is entitled to one $K(t+1)^{th}$ of physical capital's payoff in $t+1$ and each of the $N_m(t+1)$ human capital securities to one $N_m(t+1)^{th}$ of labor's payoff. The aggregated factor income of physical capital and human labor they pay each period t is:[41]

$$\tilde{\theta}(t)\tilde{Y}(t) + \frac{1-\vartheta}{\vartheta}\tilde{I}(t) \qquad \text{and} \qquad \left(1-\tilde{\theta}(t)\right)\tilde{Y}(t). \tag{5.4.1}$$

The $\tilde{\theta}(t)$-based split is evident; physical capital's share is just supplemented by the value generated by the adjustment costs. As Sect. 5.2.3 explained, when human capital is tradable, the middle-aged generation is not limited to realize investments in physical capital only. And the young cohort is not forced to be the only holder of human capital securities. Thus in the subsequent period, physical capital's share of income is not paid only to the old, and the middle-aged are not the only ones receiving labor's share of income. Instead, these generations share the "*factor income risk*" and each becomes exposed to both factor incomes. This can be seen formally by combining (3.5.2) and (5.2.11) to

$$S_y(t) + S_m(t) = S(t) = S_H(t) + S_K(t). \tag{5.4.2}$$

Young and middle-aged are not constraint to a particular form of savings in order to transfer their wealth intertemporally. The complete capital market allows the two generations to pool the risks and share the holdings of the human and real capital stocks in aggregate savings.
The resulting risk sharing does not only reduce each generation's exposure towards $\tilde{\theta}(t)$, it reduces the $\tilde{\theta}(t)$-risk entirely. To see this, one only needs to compare the return relations: in equilibrium the return on physical capital, given by (3.5.23)

$$\tilde{R}_K(t+1) = \left[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)\right]\frac{\tilde{Y}(t+1)}{I(t)},$$

and the return on human capital, given by (3.5.24)

$$\tilde{R}_H(t+1) = \left(1-\tilde{\theta}(t+1)\right)\frac{(1-\vartheta)\Psi + \vartheta\bar{\theta}}{1-\bar{\theta}}\frac{\tilde{Y}(t+1)}{I(t)},$$

are differently affected by the future realization of $\tilde{\theta}(t+1)$. The return on the *complete* capital market as determined in (3.5.22),

$$\tilde{R}_M(t+1) = [(1-\vartheta)\Psi + \vartheta\bar{\theta}]\frac{\tilde{Y}(t+1)}{I(t)},$$

[41]Remember that according to (3.2.6) and (3.2.13) the derivation of the securities' payoffs has been based on this argument. Furthermore the two components clearly add up to gross domestic product as given in (3.3.4).

however, is only exposed indirectly. The indirect channel of real capital's production elasticity is based on the creation of $\tilde{Y}(t+1)$ according to (3.2.1). Ceteris paribus, this future output will be the bigger the higher the productivity of the relatively abundant production factor. That is, when there is a relatively large stock of real capital, output will be higher if its productivity is high due to a high realization of $\tilde{\theta}(t+1)$ – and vice versa. Yet, this indirect influence of $\tilde{\theta}(t+1)$ affects all three returns in the same way via $\tilde{Y}(t+1)$! But only human and physical capital's return have additional factors that also depend on $\tilde{\theta}(t+1)$. And this dependence is not the same for the two kind of securities, but comes in a from that cancels out when all securities are aggregated to the capital market. Apart from the indirect channel via $\tilde{Y}(t+1)$, the production elasticity enters the market's return only by its mean $\bar{\theta}$. In this sense this distribution part of $\tilde{\theta}(t)$-risk that manifests itself as factor share risk is *diversifiable* under complete markets! It relates to the risk how the generated output is distributed to the input factors.
When human capital is not tradable, it has to hold that $S_y(t) = S_H(t)$ and $S_m(t) = S_K(t)$. The young's savings allocation is forced to too much, i.e. only, human capital and the middle aged have too little, i.e. no, holdings of human capital in their portfolio. The pooling of factor income risks as formulated in (5.4.2) does not take place. From an aggregate perspective this means the Overlapping Generations Separation, as introduced in Sect. 5.2.3 fails. However, the PAYGO system together with the consumption tax as calibrated in Chap. 4 heal the capital market's incompleteness also from the perspective of risk sharing: by eliminating the unnecessary – since diversifiable – factor share risk it provides the same, efficient risk sharing as the hypothetical human capital securities. The transfers from workers to current retirees eliminates the asymmetrical exposures towards $\tilde{\theta}(t)$. This interpretation of PAYGO as a mean for efficient risk sharing is the same as in Merton [1983] but with accumulation of real capital.

Undiversifiable Risks

Yet, there are also risks that cannot be mitigated by the pay-as-you-pension scheme and have not been explicitly addressed by Merton.[42] All the returns above are tied to future uncertain output. This $\tilde{Y}(t+1)$ does in its formation not only depend on $\tilde{\theta}(t+1)$ as just described, but also crucially on the future realization of the total factor productivity, $\tilde{A}(t+1)$. Each period the demographics (3.1.2) and the capital accumulation technology (3.2.7) completely determine the amount of input factors available for production in the subsequent period. However, the result of this production is only determined once the realization of $\tilde{\theta}(t)$ and $\tilde{A}(t+1)$ are revealed. Only when the factors' individual productivity – by $\tilde{\theta}(t)$ – and the total productivity – by $\tilde{A}(t+1)$ – have been identified, the relative usefulness of each input factor is known and

[42]Probably this has been too obvious, but this distinction is crucial for assessing the potential of PAYGO pension schemes.

results in the corresponding size of output. This kind of "*endowment risk*" is *undiversifiable* even with tradable human capital!

From a formal point of view it is undiversifiable because the return on both kind of securities, $\tilde{R}_K(t+1)$ and $\tilde{R}_H(t+1)$, depend in the same manner on $\tilde{Y}(t+1)$ and thus on the total factor productivity and on the production elasticity. As the impact is absolutely the same, this $\tilde{Y}(t+1)$-dependence cannot be eliminated by forming portfolios. Hence, it still appears in the return of the capital market $\tilde{R}_M(t+1)$ and also determines the consumption levels in equilibrium as seen in (3.6.25) to (3.6.27). The little additional model of Sect. 5.3 has indicated that this would also hold, when a more advanced method of explicitly addressing exogenous technological change is used. It has been shown that uncertainties in population and technology enter national income to the same extent as they infiltrate the individual factor incomes. As each of them fully mirrors the total endowment risk, this risk is not diversifiable by any combination of the factor incomes. In other words, explicit technological progress will not help to diversify it.

From an economic perspective this is like a Lucas [1978] endowment economy on a meta-level: instead of bearing $\tilde{Y}(t)$-fruits directly, the tree now generates a level of $\tilde{A}(t)$ and determines the value of $\tilde{\theta}(t)$. Based on the macroeconomic production function, these parameter-fruits allow then to generate consumable units with the endogenously determined labor force and real capital stock. As clearly seen in (3.2.1), an increasing sequence of $\tilde{A}(t)$s, interpreted as advances in the economy's technological level, will allow for the production of ever higher output. In contrast to this, real capital's production elasticity is bound by (3.2.2) to the limits of $0 \leq \tilde{\theta}(t) \leq 1$. And variations in it can cause higher or lower levels of $\tilde{Y}(t)$, depending on the relative sizes of $K(t)$ and $L(t)$. But it still makes aggregate output risky. The uncertain influences of $\tilde{A}(t)$ and $\tilde{\theta}(t)$ on $\tilde{Y}(t)$ are thus referred to as *endowment risk*. The last paragraph has established the uncertainty of distributing the output to the two input factors as *factor share risk*. As this uncertainty is driven by $\tilde{\theta}(t)$ as well, it would be ambiguous to classify $\tilde{A}(t)$-risk as undiversifiable and $\tilde{\theta}(t)$-risk as diversifiable. Only the distributional part of $\tilde{\theta}(t)$-risk can be eliminated with tradable human capital or with the PAYGO replication.[43]

Before turning to a financial interpretation, an intuitive illustration of this finding is useful. The total pie in Fig. 5.3 represents the total output generated by the economy in a single period.

Undiversifiable endowment risk is reflected in varying sizes of this pie. With advantageous realizations of $\tilde{A}(t)$ and $\tilde{\theta}(t)$ this pie is bigger, with less advantageous realizations smaller as indicted by the dotted line. Factor share risk manifests itself in the relative size of the slices. Since each factor is paid its marginal product, physical capital captures a $\tilde{\theta}(t)$-portion of the entire pie

[43]This confirms the results of von Weizäcker [1999], who lacks the $\tilde{\theta}(t)$-uncertainty and finds thus no diversification potential between the pension pillars.

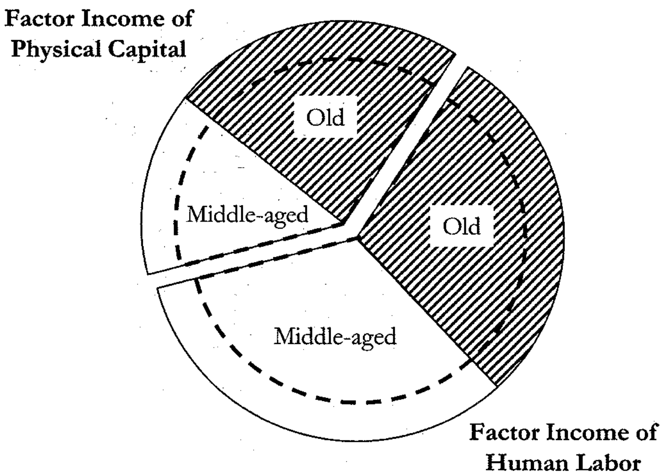

Fig. 5.3. Risk sharing

Remarks: Undiversifiable endowment risk is reflected by the risky total size of the output-pie. Due to the tradability of human capital – or the replication with PAYGO – middle-aged and old can share the factor income risk. This is the risk of how the pie is distributed to physical capital and human labor.

and human labor the remaining $1-\tilde{\theta}(t)$-portion.[44] Section 5.2.3's accounting analogy has illustrated that only the young and the middle-aged can claim this output, since only they work or own the real capital stock. When human capital is not tradable, human labor's slice would only belong to the working middle-aged and the factor income of physical capital would be paid entirely to the old generation. Both generations would then be fully exposed to the factor share risk: though the total pie might be bigger one's own slice of it might be smaller in absolute terms because of the realization of $\tilde{\theta}(t)$. But with tradable human capital this exposure is unnecessary and the two generations can share their *idiosyncratic* risks. By pooling their savings and diversifying into both kinds of securities they not only share the asymmetrical exposures but also eliminate the factor share risk entirely. In Fig. 5.3 this is indicated by the fact that middle-aged and old receive a part of each input factor's slice. This results in a total distribution of the output-pie not based on the

[44]Ideally, one would like to create this illustration with gross domestic product. Based on (5.2.20) $\tilde{Y}^{GDP}(t)$ is proportional to $\tilde{Y}(t)$ and thus inherits its risk characteristics. Yet, the value caused by the adjustment costs is only captured by physical capital as can be seen in the additive, $\tilde{\theta}(t+1)$-independent term in (3.5.23). Thus, only $\tilde{Y}(t)$ is split according to $\tilde{\theta}(t)$ but not $\tilde{Y}^{GDP}(t)$. Still, the effects of endowment and factor share risk can be explained by the same intuition.

stochastic realization of real capital's production elasticity but purely on the respective previous amount of saving. In this sense, tradability allows to eliminate the idiosyncratic exposures to factor share risk.

The pay-as-you-go pension scheme can achieve risk sharing, too, because the contributions are paid from human labor's share and give thus the old a share of labor's factor income. And the calibrated fiscal system makes a participation of the middle-aged in physical capital's factor income possible. However, neither the benchmark model with the complete capital market nor the replication with PAYGO can diversify or eliminate the endowment risk: when the output pie is bigger there is more to share – between factors or between generations. As agents do have to bear only this aggregate risk, the allocation – or sharing – of risk is *efficient* as stated by the mutuality principle of Wilson [1968].[45]

Financial Interpretations

With a decomposition of the total risk in an undiversifiable and a diversifiable part the analogy to the Capital asset Pricing Model of Sharpe [1964], Lintner [1965], Mossin [1966], and – due the lack of an riskless asset in net supply – in particular to the version of Black [1972] is evident. The general equilibrium framework developed here with its complete capital market extends the usual understanding of the market portfolio to explicitly include the capitalized value of human capital. Hence, there is no need to adjust for nonmarketable components as in the early attempts to address human capital, like Mayers [1973] or Fama and Schwert [1977]. The factor share risk inherent in each kind of security can be eliminated by diversification into the other kind. It can be considered as *unsystematic* risk. Endowment risk, on the other hand, must even be taken with a perfectly diversified portfolio and is thus *systematic*. The illustration of Fig. 5.3 provides an intuitive explanation of this distinction: idiosyncratic characteristics of human labor's or physical capital's share of output are perfect negatively correlated and thus cancel out in the market portfolio. In aggregate, only the risk stemming from the uncertain size of the output pie matters.

Finally, the diversification effect and the capital market completeness can also be illustrated in the intertemporal perspective of Sect. 5.2.3. Figure 5.4 visualizes the returns offered by the capital market for different realizations of $\tilde{A}(t+1)$ and $\tilde{\theta}(t+1)$. To simplify the illustration it is based on the case of full depreciation and on the assumption that the absolute values of $L(t+1)$ and $K(t+1)$ are equal. In this case, due to the specification of (3.2.1) the realization of $\tilde{\theta}(t+1)$ does *not* have any influence on the amount of output and determines the distribution only. In other words, endowment risk is only caused by $\tilde{A}(t+1)$.

Panel (a) focuses on this undiversifiable endowment risk: even the complete

[45] See also Gollier [2001, pp. 309-311] and Lengwiler [2004, p. 107].

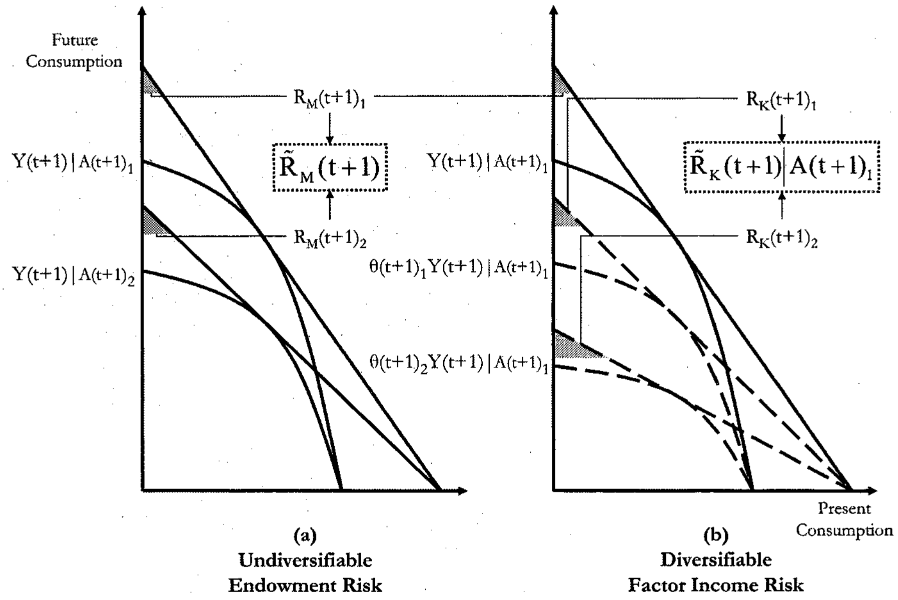

Fig. 5.4. Risk Decomposition

Remarks: Endowment risk caused by different realizations, $A(t+1)_1 > A(t+1)_2$, of the total factor productivity cannot be eliminated with the complete capital market. However, investing only in real capital implies diversifiable factor income risk because unnecessary exposure to realizations of the production elasticity parameter, $\theta(t+1)_1 > \theta(t+1)_2$, can be diversified in the complete capital market.

capital market's return $\tilde{R}_M(t+1)$ is risky since different realizations of the total factor productivity, $A(t+1)_1 > A(t+1)_2$, result in different intertemporal terms-of-trade offered by the economy, i.e. different return retaliations of the market, $R_M(t+1)_1$ or $R_M(t+1)_2$. Even with perfectly diversified portfolios agents will have a worse trade-off of current versus future consumption for $A(t+1)_2$ then for $A(t+1)_1$ because they cause different general performances of the economy.

Panel (b), on the other hand, depicts the different return realizations on physical capital for a given future total factor productivity $A(t+1)_1$. This translates into $R_M(t+1)_1$ as return on the market. Physical capital, however, captures only a fraction of this. The size of its share is determined by the realization of real capital's production elasticity.[46] This is illustrated by the payoff generat-

[46] As the adjustment costs are not present in the case of full depreciation, the additive term reflecting the value of capital per se is not present in real capital's return. See (3.5.35) in case of doubts. Yet, principally the same intuition also holds for the general case.

ing the return. When its production elasticity is higher, $\theta(t+1)_1 > \theta(t+1)_2$,[47] physical capital receives a larger factor income and there is thus a higher return $\tilde{R}_K(t+1)_1$ for holding it than for a low production elasticity $\tilde{R}_K(t+1)_2$. For the return on human capital this would be the other way around, as labor's and physical capital's share add up to total output. For this reason, factor income risk is diversifiable with a complete capital market: agents are not limited to invest in real capital only. Instead of receiving solely one of the $\tilde{R}_K(t+1)$-realizations, they can diversify this unnecessary exposure and obtain $\tilde{R}_M(t+1)_1$. As clearly seen by the steeper slope in Fig. 5.4 (b) this return dominates each return on the physical capital for a given total factor productivity and offers thus a better intertemporal trade-off.
Again PAYGO has the same efficiency-improving effect as a complete capital market: it effectively allows the agents to act as if they had $\tilde{R}_M(t+1)$ determining their intertemporal constraint. However, neither the pay-as-you-go system nor the complete hypothetical human capital securities allow to eliminate the endowment risk.

5.4.2 Earning Points and Indexation

The illustrated efficiency-improving diversification effect of PAYGO concerning factor income risk has so far been based on the assumption that all members of a cohort work when they are in their middle-age. Earning points as briefly explained in Sect. 2.2.6 would thus not have a special a role because all agents supply the identical amount of labor and are thus entitled the same pension benefits.

Elastic Labor Supply

The contribution of Merton [1983], however, addresses this important feature of actual pay-as-you-go schemes and shows how it is required in order to implement the proposed replication idea with *elastic* labor supply. Since an extension of this aspect into the more general framework of Chaps. 3 and 4 would be beyond the scope of this analysis, this section will explain how the earnings points work in Merton's [1983] case. Section 4.4.5 has shown that the Merton framework is a special case of the full depreciation model, so there should be some way to generalize his results. In order to make the labor-leisure decision endogenous Merton included an additional term in the utility functions as given in (4.4.33) and (4.4.34):

$$U_y(t)|_{\text{Merton}} = \ln c_y(t) + \mathbb{E}_t[\Gamma \times \ln l(t+1) + \ln c_m(t+1) + \ln c_o(t+2)]$$
$$U_m(t)|_{\text{Merton}} = \Gamma \times \ln l(t) + \ln c_m(t) + \mathbb{E}_t[\ln c_o(t+1)]$$

For linking the Merton-case into this framework the assumption has been that the scaling factor for utils from leisure, Γ, is zero. This has eliminated

[47] Remember, that the assumed equal numerical values of $L(t+1)$ and $K(t+1)$ prevent $\tilde{\theta}(t+1)$ from having an endowment influence on $\tilde{Y}(t+1)$ or $\tilde{R}_M(t+1)$.

any influence of $l(t)$, the fraction of an agent's work period spent in leisure, on overall utility. However, it also implied that all middle-aged work the entire time they can work. Yet, Merton [1983] also allows for $\Gamma > 0$ and thus agents to trade off leisure and consumption, making the supply of human labor endogenous. In the benchmark model it is then analytically convenient to assume that the young issue their human capital securities at a price of the *gross* value of human capital. This price is defined as the market price for the wage income the agents would earn if they were to work 100% of the time during the middle-age of their life-cycle. Based on this convention it is logical that an agent must buy back the leisure time he chooses to consume when he is middle-aged. Due to this additional expenditure savings of a middle-aged are then given by his wealth less his consumption and less his purchase of leisure time. This additional element has severe effects on the consumption-savings policy and the asset allocation. But the optimum portfolios are still *identical* for the young and the middle-aged cohort. However, with human capital securities being a claim on gross human capital, the payoff of them is not only the wage rate but also the value of leisure time purchased in the corresponding period. To some extent this is similar to the value of physical capital per se based on the adjustment costs: human capital has an extra value because the time endowment also allows to trade it in for leisure.[48] For this reason the capital market's portfolio weights in equilibrium also depend on Γ. The leisure-driven changes in individual consumption and savings are also present in the aggregate equilibrium quantities of output, consumption and investments. Furthermore, the elastic labor supply manifests itself in the fact that the supply of human labor is now determined by the size of the middle-aged cohort less the aggregated demand for leisure.[49]

This aggregate demand for leisure is an important additional condition to be met when the first-best allocations are replicated with the pay-as-you-go scheme and tax system. Without changes, the replication with PAYGO would generate a smaller equilibrium quantity of labor because the PAYGO contributions distort the labor-leisure decision in the direction of demanding more leisure. They are a disincentive to work. [50] In order to reduce the magnitude of this distortion Merton adds a third aspect of the public sector to the replication: eligibility requirements that make retirement benefits progressively depending on the relative amount of contributions made by the individual agent. He shows that a system which distributes the aggregate pension bene-

[48]Note that as Merton [1983, p. 331] remarks one could alternatively define the price of human capital securities based on the net value of human capital; then they would not yield this extra payoff. The applied definition is without consequences for the equilibrium result and purely for convenience to apply the usual accumulation equation.

[49]See Merton [1983] equation (40).

[50]In contrast to this the consumption tax has been applied to *any* kind of consumption and thus they distortion is the same along the life-cycle and therefore without consequences for the case of *myopic* log-utility.

fits to the old-aged agents according to their individual contributions relative to the average contribution, can indeed eliminate the distortion caused by leisure. With such a system of eligibility requirements a worker, when determining his optimal quantity of leisure time, will not only take into account the loss of income, but also the loss of retirement benefits. This makes leisure more expensive and – when carefully calibrated – induces the same labor supply as without the PAYGO contributions. These eligibility requirements can be compared to the system of earnings points or average indexed monthly earnings encountered in actual pay-as-you-go schemes.[51] From this perspective, earnings point are less needed to account for the implicit PAYGO savings than as an incentive to work.

Furthermore, as pointed out by Merton [1983, p. 348], the PAYGO scheme cannot be a voluntary system. Though it has been shown in the replication framework that each period exactly the efficient complete markets' allocations are realized, a single cohort could have an incentive not to participate. This might happen when the current realization of $\tilde{\theta}(t)$ yields a very high wage rate but the typically observed values imply that this cohort, once retired, is likely to receive a much smaller total PAYGO budget. Then the equilibrium shadow present value of aggregate retirement benefits to the current workers would be less than the aggregate PAYGO contributions paid by them. In that case, they would not stay voluntarily in the system, unless participation was still desirable in order to share risks as explained in the previous section. This highlights the insurance character of pay-as-you-go schemes: participation might be beneficial even though contributions exceed the shadow present value of benefits because PAYGO serves as a perfect hedge against negative – regarding its impact on physical capital's factor income – fluctuations in the production elasticity. Note, however, that this would require to go beyond the simple assumption of independent and identically distributed $\tilde{\theta}(t)$s made in Sect. 3.2.1. Still, the system of earnings points can also be seen as a mean to supervise the mandatory participation in the PAYGO scheme.

Indexation

Having explained the role of earnings points from the perspective of replicating human capital tradability, one can also address the typically encountered indexation of pension benefits. The stylized description here will follow mostly Börsch-Supan and Wilke's [2003] and Borgmann's [2005] analysis of the German pension system, but since other countries have very similar mechanisms the results are not limited to the German case.

While earnings points reflects a retiree's claim relative to his cohort peers,

[51] See, for instance, Börsch-Supan and Wilke [2003]. Furthermore, overcoming the strict life-cycle of this model and allowing for individually chosen retirement ages requires to complement the earnings point systems with the concept of years of service life.

the indexation is used to determine the *current pension value*[52] as reference point for this claim. In other words, the earnings points drive the intra-cohort distribution and the indexation the intergenerational distribution. Typically, the indexation is also seen as a descriptor of PAYGO's generosity. The reason for this is that actual pension schemes do typically *not* operate under the strict budget rule of (4.2.19). Instead, benefits and contributions are determined separately by law. Current rules in most countries do not require that benefits and contributions equal, but allow for the accumulation of deficits by operating at politically desired low contribution rates and grant at the same time high replacement rates. The cohort-average replacement rate is crucially driven by the development of the current pension value. A specific pension benefit indexation formula links this value to present variables of the economy. Even though the indexation formulas have changed several times during the last decades, across countries they have the characteristic feature of linking pension benefits to the level of wages and salaries. For Germany, for instance pensions were indexed to gross wages until 1992, to net wages until 1998 – then two years to inflation – and since 2001 a rather complex new adjustment formula has been effective, which relates changes in the pension value to lagged changes in gross income.[53]

Based on this framework, this linkage of pension benefits to the wage *level* is seriously inappropriate for addressing the problem of increasing old-age dependency ratios and shrinking labor forces as described in Chap. 1! It is flawed because it neglects the underlying economic changes when the population declines: imagine a situation in which the working middle-aged generation consists of only very few agents. Because of the relatively high stock of physical capital per worker these would be extremely productive and hence earn a very high wage rate. This balancing effect is inherently caused by the neoclassical production function and input factors being paid their marginal product. The compensation is based on human labor's and physical capital's relative contribution to the production process, so that relatively scarcity leads to a higher compensation *per unit*. But in total each factor captures only a portion of output determined by his aggregate productiveness. This is the underlying principle of Chaps. 3 and 4.

When in the imagined situation of very few workers the pension benefits of the currently old are adjusted to the high wage level due to some indexation, the PAYGO system will face the choice of either running deficits or increasing the contribution rate. In the language of Fig. 5.3, labor's slice of the output pie is not increased – it is just distributed to fewer workers making the individual agent's slice, i.e. his wage, higher. In order to generate sufficient funds for financing the increasing expenditures to retirees the PAYGO system can

[52]See Börsch-Supan and Wilke [2003, p. 12].

[53]See Borgmann [2005, Table 6.3] for a good overview and detailed description on Germany. In the United States the system is based on averaged indexed monthly earnings; see Rosen [2002, pp. 181-185] or Liebman [2001].

only command a higher fraction of labor's share, that is a higher contribution rate – or run a deficit by issuing explicit or implicit debt. While the aim of the indexation is to enable pensioners to share the rising prosperity generated by the economy, its implementation based on the *level* of wages is not neutral regarding demographic changes. The problem is that the wage level reflects both influences: those of a rising prosperity due to the accumulation of physical capital or advances in the total factor productivity as well as those of distributing labor's share of output to individual workers. Wages are like other prices indicators of relative scarcity in the economy. Yet, scarcity of human labor is driven by two separate forces: demographic change in course of global aging and technological progress giving individual labor units a higher economic value. It might well be that a society chooses to let retirees generously participate in the latter, yet the former reason for scarcity does not allow for generosity. It merely reflects the rising problem of shortage of human beings and is thus the opposite of rising prosperity! For an increasing population, as temporally caused by the post war Baby Boom, the wage indexation has the advantage that pension benefits rise slower than the economy grows because the additional supply of human labor depresses the wage level. But for a declining population – characterized by ever viewer workers per retiree – it aggravates the pension problem by overstating human capital's actual factor income. In other words, for the pay-as-you-go system not the individual levels matter, but aggregate equilibrium amounts.

The obvious consequence is that replacement rates are also the wrong mechanism for governing PAYGO systems. They are merely the result of it, just like realized returns for stock market investments. When there are only relatively few workers, there is no economic way to safely guarantee the non-working old generation pension benefits at some 75% of their previous labor income. The few middle-aged can only generate a smaller output pie than their predecessors unless they receive a very favorable parameter-endowment of the meta-Lucas [1978] tree: a very high $\tilde{\theta}(t)$ would make human labor less important to the production process and a high $\tilde{A}(t)$ would increase the output directly. And, as emphasized before, only what is produced is distributable – among the production factor or between the generations. In this sense, the Mackenroth [1952, 1957]-hypothesis mentioned in Sect. 2.2.6 also holds in the portfolio perspective: all retirement expenditures have to be made from the current national income, regardless how the combination of funded and unfunded pension systems looks like. The diversification across financing pillars only allows to eliminate the *additional* risk in the functional distribution of national income to the macroeconomic input factors.

In reality the indexation resulted in financing problems of actual PAYGO system and puts increasingly pressure on governments to change the pension formulas. In Germany for instance, a reform commission for "sustainability in financing the social insurance systems" was established. The recommendation of the popularly called "Rürup commission" does not only include further cuts in the replacement rate of the pay-as-you-go pillar, but also a change in the indexation formula for the current pension value: it proposed the introduction of an additional factor that leads to higher pension benefits in case of an increase in the number of contribution payers and to lower pensions in case of more beneficiaries. The "Pension Reform Sustainability Act" realized these recommendations.[54] By loosening the tight dependence of benefits on wage levels and linking pension adjustments to the crucial factors determining pension financing, namely the number of contributors and benefit recipients, this change gives the new pension benefit indexation formula to some degree a self-stabilizing effect. The "Rürup-sustainability" factor does theoretically allow that a pension increase driven by a higher wage level is overcompensated by a deteriorated demographic situation causing the wage increases. However, grandfathering rules guaranteeing certain minimum levels prevent an actual overcompensation. Still, the changes in the formula index pension benefits less to the problematic wage *level* and more to human labor' aggregate *share of output*. Yet, it seems that the determination of pension benefits is still too much focussed on individual levels instead of fixing the underlying problem by completely switching to an indexation to labor's total factor income.[55]

5.4.3 Replication Framework and Existing PAYGO Schemes

While the last section has already indicated some problems of actual pay-as-you-go pension schemes, a more thorough comparison of them to the framework developed here is mandatory. Furthermore, the results of the replication framework require some qualifying assessment.
As mentioned in Sect. 4.4.3 the replication solution resulted in extremely high contribution rates. As seen in Fig. 4.2, τ_ω seems to exceed 60% for all combinations of the parameters ϑ' and $\bar{\theta}'$. Actually existing pay-as-you-go system, on the other hand, have contribution rates of only around 20%. The resulting quantitative gap makes not only benchmark exercises of actual PAYGO schemes useless, but might also cause doubts on the proposed framework, the derived implications and their significance for reality. There are two sources causing the gap: the simplicity of the model, but also the shortcomings of real pay-as-you-go pension schemes.

[54] See Bundesministerium für Gesundheit und Soziale Sicherung [2003, p. 103]. Note that already the later revoked reform of 1999 would have made such an adjustment: the effect of Rürup's sustainability-factor is essentially the same as that of the early proposed demographic-factor. Admittedly, it must also be mentioned that the financing problems are partially caused by structural problems on the labor market resulting in economic and unemployment sclerosis. Yet, the transmission mechanism for the PAYGO problems is the same: less workers and a high wage level.

[55] Admittedly, indexation of pensions to wages has also an important role in the presence of monetary issues: as wages adjust for inflation so do the pension entitlements. While this monetary aspect cannot be addressed in this framework in real terms, one could expect that direct links to inflation could better accomplish the objective of inflation adjustments.

Confronting Reality

Starting with the former reason, the framework proposed in Chaps. 3 and 4 is overly simplistic. As introduced in Sect. 3.1.1 there is an overlapping generations model underlying, which consists of only *three* cohorts: a growing up young one, a working middle-aged one and a retired old cohort. There is thus no distinction concerning agents in their early and late working years. Neither is the consumption life-cycle addressed beyond the implications of the three phases. There are no durable goods, nor is real estate addressed as an important consumption medium or investment category. Bequests have been ruled out. Risk aversion is assumed to be constant throughout the life-cycle, even though increases towards the end of life is highly intuitive. The economy is closed, so that neither the goods market can profit from international trade nor can investors diversify their portfolios internationally. And the replicating public sector with its PAYGO and fiscal system is very stylized. All these simplifying assumptions have been made in order to keep the model analytically tractable and to focus on the efficiency argument. Yet, they all depress the investment-output ratio Ψ as the most important descriptor of the economy. According to Fig. 3.3 its maximum values are only around 10% which is rather low compared to typical observed numbers of about 25%.
But with a period length of around 30 years only those expenditures qualify as investments that increase the physical capital stock of the subsequent period – i.e. 30 years later – making the 10% less unrealistic. The neoclassical paradigm extremely reduces the firms' role: all output is sold and revenues are completely distributed to the input factors, leaving no economic profits. There is not even a distinction between different firms or sectors. Yet, typically expenditures for productive machinery amounts only to roughly 40% of investments, while 60% are spent in the building and construction sector, which are not explicitly addressed here. Openness of the economy would not only allow for international trading in the output market, but also for an international dimension of the risk-sharing discussed in Sect. 5.4.1.[56] With additional diversification domestic investment in physical capital will be more attractive and thus increase Ψ. The same should result from more general utility specifications and more frequent adjustment in the policies. Dynamic asset allocations rules with more than the one theoretical rebalancing will make investments of any kind less risky since adjustments to stochastic changes can be implemented sooner and more often. This will favor future consumption at the expense of present one and thus increase national savings. To sum up, the framework's assumptions have resulted in a comparably low investment ratio, but when lifted will drive it in the right direction.
However, the simplifications have not been made deliberately, but in order to show *analytically* that the replication does work. Generalized versions will

[56]There is even an asset management strategy dedicated to this argument; see, for instance, Cragg [1998].

certainly require numerical approaches to achieve this goal. But the accompanying inaccuracies are hardly of negligible order in case of macroeconomic long-run simulations! Furthermore, even though the investment-output ratio hardly matches realistic values, the final distribution of output to human labor and physical capital mirrors the results of growth theory rather well: with real capital's production elasticity fluctuating around 0.3, two thirds of output are expectedly paid to labor and one third to physical capital. And this is what ultimately determines the relative attractiveness of investments in both kind of securities. But this distribution is also important for the functioning of pay-as-you-go pension schemes: such systems derive their financing from the total amount paid in form of wages and salaries. And as Sect. 5.2.1 has emphasized again, human capital is nothing else than a present value statement of future wages. This key insight has been the cornerstone of the replication idea. Yet, if human capital captures around two thirds of output and the simplifying assumptions imply that middle-aged and old have the same portfolio exposures to each input factor, it is *not* surprising that the replications requires a contribution rate of more than 50% to achieve this. Since the old must co-finance the consumption tax, they command an even higher fraction of labor's income, in order not to be made worse off.

Generational Accounting

In addition to this assumption-driven explanations, it is also noteworthy that the framework goes beyond the traditional economic income accounting. With the embedding of national wealth as given in Sect. 5.2.3 it realizes a step towards generational accounting:[57] by using present values and not current flows it prevents from income accounting's pitfalls to split continuous actions in periodic reports. This is comparable to the short-term orientation managers or corporations are frequently accused of: instead of acting in favor of the company's long-run performance they might focus on those goals only which have a prompt impact on quarterly figures. The same basic problem of accounting seems to cause the problems on the national or on the corporate level: it is quite easy to account for changes in revenues because they occur in the course of a specific period. But it is a more complicated task to assign the underlying causes and thus to define periodic profits and prosperity properly. For companies it might be superior long-run product development or short term price reductions that cause high revenues, for economies it might be slow demographic changes, unexpected productivity alterations or cyclic behavior throughout the business cycle that influence national income.
The great virtue of the intertemporal perspective is not the proposition that such future changes can be anticipated by financial markets, but that the uncertainty accompanying this anticipation can be incorporated properly. The

[57]See Auerbach et al. [1991, 1992, 1994] for the academic development of generational accounting. An international overview of its application is provided in Auerbach, Kotlikoff and Leibfritz [1999]. The European Commission [1999] and Raffelhüschen [1999] focus on the European Union.

concept of wealth – on an individual or aggregated level – overcomes the shortcoming of income: it does not deliberately assign anything to a single period, but tries to assess and account for all future changes of a current action. By addressing future consumption possibilities it accredits today additional flexibility. Unfortunately wealth accounting has been seen only as a supplement to the traditional national income accounting framework. In particular the inclusion of human capital – as the most important part of it – proofs to be quite complicated. While the simplistic framework here allowed to derive closed form solutions for human capital easily, it is a different task to assess the capitalized value of future wages when wage levels have to be estimated over a varying working life of around 30 years – not to mention cross-sectional differences due to human capital formation by education, experience or training. But in addition to this measuring problem, human capital might have also been neglected rightfully because shortage of human labor has hardly been an issue in the past.[58] Yet, given the future demographic projections of Fig. 1.1 it will certainly be in the future.
Ignoring the externality of wars, history of mankind has been characterized by population increases, in particular since the Agricultural Revolution as Chap. 1 illustrated. But now this enormous long-term trend will change, ultimately leading to shrinking populations and thus to a shift from physical capital to a focus on cohorts and human capital. This will also reduce the role of periodic income accounting and emphasize the intertemporal perspective of generational accounting. As intuitively shown in Kotlikoff and Burns [2004], it is this typically neglected constraint of the public sector that matters. It is easy to generously make pension promises, when the resulting liabilities do not have to be accounted for because they do not affect *current* output or the present government budget. Yet, the burden a future generations is faced with *does* include this implicit debt:[59]

Present value		Present value		Official		Implicit		Present value
of net taxes of	=	of government	+	public	+	public	−	of net taxes of
future generations		purchases		debt		debt		current generations

In the framework of Chaps. 3 and 4 official as well as implicit debt is zero: the public sector's budgets are always balanced and the old's PAYGO benefits are not based on some past promises which would generate unaccounted implicit liabilities, but purely on human capital's share of *realized* output. The replication does not distribute an output pie which has not yet been produced. This is a kind of extension of the Mackenroth [1952, 1957]-hypothesis to the case with uncertainty and accumulation of physical capital.

[58]Note, that this refers to the perspective of a single country or of the entire world. Important contributions to the understanding of different qualities of labor input have been derived in inter-country studies; see for instance Uzawa [1965], Lucas [1988] or Lucas [2002].

[59]See Kotlikoff and Burns [2004, p. 49]

Existing PAYGO Schemes

Existing pay-as-you-go schemes, however, are typically not characterized by generational neutrality. And this shortcoming is the second reason for the quantitative mismatch in the contribution rates. Existing PAYGO systems accumulate *immense* implicit debts: for the United States Kotlikoff and Burns [2004, pp. 51] estimate an implicit debt from Medicare, Medicaid and Social Security of $43 trillion, i.e. ten times official debt. For Germany Fehr and Habermann [2004, Table 3] estimate an sustainability gap of up to 200% of gross domestic product in the worst case – more then three times official debt. In other words, there is *no* way that PAYGO systems can remain as they are currently implemented: they are deeply unsustainable and past and proposed reforms are only superficial. Therefore, it would be unfair to benchmark the τ_ω-levels from the replication framework against the now observed contribution rates of actual PAYGO systems.
The perspective of efficient risk sharing as presented by this framework has hardly been introduced into the public discussion. Instead, fairness arguments dominate the public awareness. Furthermore, Sect. 1.2 explained that pension system serve also a politically desired redistributive function: widows receive benefits even though they have not paid contributions, motherhood receives earning points, unemployment problems are shifted into the pension system by actuarial unjustified retirement rules, government employees or self-employed are exempt from the system or have their own, etc. All this redistributive interferences may be justified from a political point of view – yet they deeply change the understanding of PAYGO: The pension system's redistribution role seems to dominate the intergenerational dimension of old-age provision. Most evident is this double-role in the fact that PAYGO system receive increasingly financing from general fiscal budgets. In Germany, for instance, the federal subsidy to the pay-as-you-go system is about 30% of the pension system's expenditures.

Synthesis

Accounting for all these differences will considerably reduce the contribution gap – the difference between the replication framework's 60% and the observed 20%. The German federal budget subsidy alone implies an actual contribution rate at around 30%. Contributions for health and long-term care which have also substantial subsidizing effect for elderly would add to this. On the other hand, generalizing the framework with more periods and generations, openness of the economy, less restrictive utility assumptions and more detailed split of investment categories should reduce the required contribution rate for replicating the complete capital market. In addition to this, the assumption of a *total* failure in the market for human capital securities placed the theoretically heaviest burden on the PAYGO system to correct the failure. In the course of a more detailed life-cycle borrowing partially against remaining human capital might well be possible, when agents can demonstrate a certain

track-record.
Furthermore, it has been emphasized in Chap. 1 that the efficiency perspective on PAYGO is a very special view and certainly not the dominating one. As the first-best allocation with a liquid market for human capital does not exist, a complete replication of its theoretical allocations might over-emphasize the diversification idea. Though it has proven as the scholastic solution to integrate PAYGO into the portfolio approach, it has never been the intention to shield retirees from relative productivity variations. Instead, pay-as-you-go schemes have been invented as the *sole* instrument of old-age provisions for entire societies. By insuring participants against some longevity risk, social planers have exposed them to the political risk of changes in the system and have placed high burdens on future generations. This contribution does by no means want to suggest that the replicating framework is the only rationale for PAYGO. However, regarding systemic design and risk diversification between different pillars of old-age provisions it should be a valuable reference point with its rigorous addressing of the diversification idea in the financing methods. The pension system's insurance function against old-age poverty has thus been explicitly neglected. While the World Bank [1994] assigned the unfunded first PAYGO pillar a prominent role in providing a minimum old-age income, the past ten years have obviously proven that this welfare function is difficult to implement with the pay-as-you-go mechanism. In the extended five pillar approach of Holzmann and Hinz [2005] the new "zero pillar" takes the difficult task to deal more explicitly with the objective of poverty prevention and to guarantee minimum living standard for the elderly. This additional system in the pension portfolio is meant to be *tax*-financed and hence rather independent form the changes in macroeconomic input factors. In this sense, the World Bank's recommendation is not so different from the portfolio approach here: redistribution shall be implemented in other means and all pillars do maintain the diversification idea.
The idea of efficient risk-sharing can even be extended to a more complex setting: if a less simplistic description of the life-cycle implies that old agents are indeed more risk averse than younger ones, then – from a risk-perspective alone – it might be rather efficient that younger generations shoulder more of undiversifiable risk than older ones. But for all pension pillars there is a limit to any risk bearing and politically desired redistributing as Sect. 5.4.1 has emphasized: only what is produced can be distributed. If fulfilling past pension promises requires more than the current output, young generations will default on the pension systems – because it is economically simply not possible to maintain them.

5.5 Summary of Implications

This chapter has provided further insights on how the replication of tradable human capital with a pay-as-you-go pension scheme works. It has been shown that the framework is well embedded in financial and economic theory and does not contradict their results as required in Chap. 2. The chapter also addressed the likely objection that technological progress should allow for some additional diversification. By briefly sketching an additional model with explicit Harrod-neutral progress it has been shown that this is not the case.
Table 5.1 compares important formal aspects of the model's first-best version in Chap. 3 and its second-best calibration in Chap. 4. The replication is initiated by the replication in the market of consumable output. Supply of goods is identical in both versions of the framework because of the same underlying Cobb-Douglas production and accumulation technologies. On the demand side, the PAYGO replication has been calibrated so that it matches the allocations of the first-best benchmark.
Savings in both version, however, differ: while savings invested in physical capital are identical, only the benchmark model allows agents to invest in human capital securities. In the PAYGO replication the capital market is incomplete so that portfolio allocations to human capital cannot be realized. Yet, agents face now a consumption tax and participate in a pay-as-you-go-pension scheme. This leads to additional elements in the intertemporal budget constraints reflecting the usual additivity problem concerning labor or transfer income. On the other hand, in the complete markets' setting the intertemporal constraints can be formalized in the standard accumulation representation since the hypothetical human capital securities allow to include labor income in the portfolio perspective.
The intertemporal connection based on gross consumption implies that wealth of the middle-aged and old generation reflects gross domestic product in either version: the macroeconomic inputs of human labor and physical capital are provided or owned by these cohorts. Hence, they own the claim on the entire valued created. In the first-best version these agents determine the young's wealth by their willingness to save and allocate savings to the human capital securities; in the second-best it is the part of wealth consumed that determines the young's wealth – via the transfer payments financed by the consumption tax. Thus in both versions national wealth exceeds gross domestic product.
With identical consumption technologies and equal savings allocated in physical capital it is logical that payoffs and return to real capital are equal. The framework's two versions differ, however, concerning the other available return: while in the benchmark case investors can directly participate in labor's share of output, the human capital securities are missing in the replication. Instead, the PAYGO system allows for an similar exposure of the old.
Therefore, the first- and the second-best version provide the same efficient risk sharing between the generations: agents only have to bear the undiversifiable endowment risk of $\tilde{A}(t)$ and $\tilde{\theta}(t)$. The distributive influence of $\tilde{\theta}(t)$ driving the

Table 5.1. Comparison of first- and second-best model

	First-best Benchmark Chap. 3	Second-best Replication Chap. 4
Goods Market Replication		
Supply Side	$\tilde{Y}(t) = \tilde{A}(t)K(t)^{\tilde{\theta}(t)}L(t)^{1-\tilde{\theta}(t)} = \tilde{Y}'(t)$	
Demand Side	$C_y(t) = C'_y(t)$ $C_m(t) = C'_m(t)$	$C_o(t) = C'_o(t)$ $I(t) = I'(t)$
Savings in		
Real Capital	$S_K(t) = p_K(t)K(t+1)$	$S'_K(t) = p'_K(t)K'(t+1)$
Human Capital	$S_H(t) = p_H(t)N_m(t+1)$	$S'_H(t) = 0$
Intertemporal Budget		
Constraint	$w_a(t) = c_a(t) + s_a(t)$	$w'_a(t) = (1+\tau_C)c'_a(t) + s'_a(t)$
Middle-aged	$W_m(t) = N_m(t)\times$ $\left[s_y(t-1)R^*_{P,y}(t)\right]$	$W'_m(t) = N'_m(t)\times$ $\left[s'_y(t-1)R'^*_{P,y}(t) + (1-\tau_\omega)\omega'(t)\right]$
Old	$W_o(t) = N_o(t)\times$ $\left[s_m(t-1)R^*_{P,m}(t)\right]$	$W'_o(t) = N'_o(t)\times$ $\left[s'_m(t-1)R'^*_{P,m}(t) + \mathfrak{t}_o(t)\right]$
Gross Domestic Product	$Y^{GDP}(t) = W_m(t) + W_o(t)$	$Y'^{GDP}(t) = W'_m(t) + W'_o(t)$
National Wealth	$W(t) = Y^{GDP}(t) + W_y(t)$	$W'(t) = Y'^{GDP}(t) + \mathfrak{T}_y(t)$
Return on		
Real Capital	$\tilde{R}_K(t+1) = \frac{\tilde{Y}(t+1)}{I(t)}\times$ $[(1-\vartheta)\Psi + \vartheta\tilde{\theta}(t+1)]$	$\tilde{R}'_K(t+1) = \frac{\tilde{Y}'(t+1)}{I'(t)}\times$ $[(1-\vartheta')\Psi' + \vartheta'\tilde{\theta}'(t+1)]$
Human Capital	$\tilde{R}_H(t+1) = \frac{\tilde{Y}(t+1)}{I(t)}\times$ $(1-\tilde{\theta}(t+1))\frac{(1-\vartheta)\Psi+\vartheta\bar{\theta}}{1-\bar{\theta}}$	n.a.
PAYGO	n.a.	$\tilde{\mathfrak{R}}(t+1) =$ $\frac{1-\tilde{\theta}'(t+1)}{1-\theta'(t)}\frac{\tilde{Y}'(t+1)}{Y'(t)}$
Risk Sharing		
Diversifiable	Factor share risk of $\tilde{\theta}(t)$	
Undiversifiable	Endowment risk of $\tilde{A}(t)$ and $\tilde{\theta}(t)$	

Remarks: The table compares the aggregate effects of the first-best benchmark from Chap. 3 and the replication with PAYGO in Chap. 4. Not only the allocations on the goods market are replicated, but also the same efficient risk sharing between generations can be achieved.

factor share risk can be eliminated by diversification – either by investing in human capital securities directly, or by participating in a wage-driven pay-as-you-go pension system. This is the central advantage of diversifying between unfunded and funded pension arrangements.

As the discussion of existing PAYGO schemes has shown, even a risk perspective cannot mitigate the underlying economics: only output that is produced can be distributed between input factors or among generations. No pension pillar can neglect the intertemporal connections in the long run. Therefore, the currently high replacement rates based on the indexation of pensions to wage levels and previously accumulated earning points are not sustainable. Hence, the gap between actually observed contribution rates and those derived from this framework are not only due to its simplicity but also due to the miss-construction of real PAYGO systems.

6

Conclusion and Further Research

Boomers are in power, but they are not solely in charge. Each of us – young, middle aged and old – is responsible for forging the path ahead. And each of us must decide whether to be part of this problem or part of its solution.

Kotlikoff and Burns [2004]: "The Coming Generational Storm", p. 247

Based on the observation of an unprecedented change in the global demographic development, the preceding chapters have proposed a complex replication framework to assess the risk diversification potential between alternative methods of pension financing. This chapter concludes with a summary on the insights into PAYGO as a mean of risk sharing, points out simplifying shortcomings of the analysis, and sketches necessary further research for advances in Pensionomics. The financial optimization within the portfolio of pension mechanisms cannot help with the economic limitation to national income as sole resource for each period's standard of living.

6.1 The Model and its Implications

The replication framework developed in Chaps. 3 and 4 and analyzed in Chap. 5 provided some new insights on the pension financing mechanisms within the multi-pillar portfolio approach. Before sketching possible further extensions, this section highlights some of the major findings.

Key Findings

Based on the dimension of the pension problem in course of global aging illustrated in Chap. 1, it has been argued in Chap. 2 that a thorough contribution in this field requires some re-integration of financial and economic theory. Consequently, this thesis has tried firstly to highlight some major theoretical advances in the understanding of financial and economic markets. Historically these have developed relatively separately, but all have some connection to the pension problem. In a second step, a general equilibrium framework of output, labor and capital markets has addressed many of these established theories in order to thoroughly integrate the consequences of the expected, unprecedented demographic change. While risk and uncertainty associated with the macroeconomic input factor of physical capital can be modeled with elaborated tools from financial theory, human labor – as the other major input

factor – is of special nature. The inseparability of labor from the individual human being prevents the application of the financial discounting methodology and its handling of risk. In order to still address the stochastic aspect of labor income, which also drives uncertainty of the pay-as-you-go pension system, an indirect implementation of the pension portfolio idea has been chosen: following the idea of a second-best solution, PAYGO has been considered as a surrogate for tradability in human capital. This resulted in a two-step-procedure of a benchmark model and its replication.
The replication framework allows for the desired portfolio-based macro-perspective on funded and unfunded pension arrangements. On the one hand, this view yields a new understanding concerning the diversification potential between the pension financing mechanisms. On the other hand, it leads to important observations concerning the design of PAYGO systems. It turns out that by consequently following the multi-pillar approach uncertainty in the distribution of national income to the factors of human labor and physical capital can be eliminated. In contrast to a situation in which old-age income relies exclusively on a single financing-method, the combination of the funded and unfunded mechanisms allows to perfectly hedge against adverse developments in the relative importance of the macroeconomic input factors. If agents derive income from both input factors, they are only exposed to undiversifiable endowment risk. In reality however, this risk sharing perspective has hardly been the motivation for combining the pension pillars. So far, financing problems in the traditional PAYGO systems have been the predominant reason for encouraging a combination of funded and unfunded pension arrangements. Consequently, the split in the pension portfolio optimized for risk sharing as derived from this framework is not observed in reality. Furthermore, this analysis has neglected any redistributional role of the PAYGO system and focussed solely on the economic efficiency argument. While there are good ethical and political justifications for the additional goals, they have typically resulted in the accumulation of huge implicit deficits as contributions and benefits are determined separately. In this respect, Chap. 5 pointed out major problems of the current indexation of benefits to the wage level. The expected demographic change will make human labor relatively scarce; wages – as any market price – will reflected this scarcity, so that they do not exclusively reflect general increases in the economy's prosperity. Benefits from an unfunded, labor-based pension system should therefore be tied to labor's total factor income and not to the individual wage rate.
The elaboration of the replication framework on basis of the financial intertemporal consumption consideration stresses also the intertemporal perspective for the public sector. As implicit form of public debt, PAYGO deficits can remain undetected in a periodic perspective. By an application of finance's standard technique of discounting, generational accounting does not only assess the immediate consequences of a public policy action, but also tries to account for all uncertain, future changes of it. If such measures were implemented thoroughly, the lack of sustainability in currently generous PAYGO

promises would be more apparent. The modeled replication framework has set the heaviest condition on the PAYGO system and completely prevented such a strategy by excluding any form of debt.

Extensions

However, with only three overlapping generations, a single sector of physical capital, no international openness and simplistic utility assumptions, the model can only present a first step towards Pensionomics. The implications clearly suffer from these shortcomings, and there are many links for extending the framework.

Further research could generalize the utility specification towards the power form and prevent the myopic implications of the logarithmic case. Habit formation, which has gained a lot of attention because of its reasonable implications and good analytical tractability, is also ideally suited for the life-cycle problem.[1] In the first-best version of the model, the key problem will then be the identification of the market portfolio in the capital market. When asset demand does not only consist of a myopic component, but also includes some Merton [1973]-like hedging demand, which probably depends on an agent's age, the portfolio allocations will no longer be the same for different generations. Hence, it matters who the marginal investors is. In the second-best version the replication of such dynamic portfolio rules will certainly require some dynamics in the PAYGO contribution rates, too.

Refining the overlapping generations model by extending it to more generations and reducing the length of a period correspondingly should yield major advances as well. So far, with only differentiating between non-working young, working middle-aged and retired old agents, the framework is limited to the simplest version able to produce the desired effects. As mentioned, this drives to some degree the quantitative implications of the risk sharing solution in the pension portfolio. Actual populations, however, are characterized by considerable more diverse age profiles than this simple distinction. Though substantially more complicated, a model with a single cohort per calendar year would be desirable in order to make the framework directly useable for calibrations with realistic population structures. Such a generalization of the life-cycle will also allow for a more realistic addressing of diverse earning-profiles of the individual life. If there is not only a single 30-year period in which an agent works, agents could be differentiated according to some labor productivity factor. This would also allow to address the important notion of human capital formation by education and training. Furthermore, some earnings early in the life-cycle could allow to eliminate the auxiliary fiscal system with the transfer payments to the young because such agents could finance their consumption

[1] This should also result in more substantial influence of the adjustment costs in the capital accumulation technology: Jermann [1998], for instance, requires habit formation in the utility specification in order for q-Theory to play a significant role.

by other means. Abandoning the strict assumption on the overlapping generations model could also go into the direction of giving more weight to the role of families. As mentioned, the intra-familiar transfer of consumable units to the offspring – as well as the wealth transfer from the deceased to the living by bequests – present theoretically important intergenerational links and can actually amount to substantial quantities. Neither has been addressed so far. However, it is doubtable whether such generalizations can still be achieved on a purely analytical level. Section 4.3 illustrated the problems of standard approaches in solving the model. The alternative solution mechanism has heavily relied on the simplicity of the generational structure and of utility specifications. With generalizations in this respect, it is hardly applicable any more. Simulations could be a viable path: one could first derive the probability distributions of the allocations in a generalized first-best version based on some distributional assumptions of the risk drivers in the economy's real side. In a second step, the public sector parameters in the incomplete markets' version of the model must be calibrated in such a way that simulations of this version replicate the first-best distribution of allocations. Clearly this would weaken the replication argument, since it would refer then only to the distributions. But possibly, even simulations of state-by-state replication are realizable.

In order to extend the framework towards different kinds of physical and human capital, one could assume different sectors within the economy's production side. When these are characterized by specific technologies and use different quantities and qualities of the macroeconomic input factors, a great variety of different physical capital and human labor classes will result. Adding furthermore some idiosyncratic risk components in the sectors' production technologies like for instance in Brock [1982], makes this aspect more similar to the standard models of the capital market like the Sharpe [1964]-Lintner [1965]-Mossin [1966] Capital Asset Pricing Model, the Intertemporal Capital Asset Pricing Model of Merton [1973], or Ross's [1976] Arbitrage Pricing Theory. The differentiation of production sectors based on the impact of demography could be especially interesting: on the one hand, there might be goods and services that are expected to profit extraordinarily from global aging, like the pharmaceutical sector, on the other hand there might be sectors which are especially negatively affected by the scarcity of human labor, like the service sector in general. Land as a fixed factor is probably less important for the production side than as an asset class yielding housing service to its owners.

Another dimension for generalizations is international openness in the economy's various markets. By sketching a multi-country or multi-region framework in different phases of the Demographic Transition, the international economic implications of global aging could be integrated. Though the aging trend is a global phenomenon, some faster greying, industrialized countries could have the temporal advantage of possessing considerable claims against other countries. By increased imports they could to some extent buffer the immediate problems of domestic labor scarcity until the foreign assets are melted down. One can also investigate the diversification potential of investments by faster aging nations in slower aging economies. Of course, such strategies will critically depend on the exchange rate development so that the unaddressed monetary dimension must also be included. The strict link to the domestic labor force can be weakened by allowing international migration. Cross country analysis with differences in the agents' preferences and the real sector's structural parameters could also shed some light on the locally most appropriate design of the PAYGO system.

But even without addressing such additional aspects, the framework can be used to enhance the understanding on the linkage of finance and economics. For instance, one could combine the results for the returns on the individual securities and on the complete capital market from Sect. 3.5.4 to derive an explicit beta-relationship as in the standard Capital Asset Pricing Model. Though the strict analogy is only achievable with the distinction of different sectors, such an extension would integrate labor as production factor and hence give some understanding on the implicit market-beta of human capital. The formulation for the riskfree return can be integrated into a normative term structure model for the real interest rate. Of course, progress in this respect requires the integration of considerable more generations so that the accompanying shorter period length allows to address maturities below 30 years. For the modeling of the PAYGO system, Sect. 5.4.2 has indicated that and how an endogenous decision on labor supply and leisure consumption could be integrated. With mean-reverting processes in the stochastic components of the production sector, one might even allow for some limited deficits of the public sector, yet caution should be taken in this respect.

These possible extensions of the model illustrate that Chaps. 3 and 4 have limited themselves to the very simplest form. This is naturally accompanied by numerous assumptions which might on their own seem – overly – simplistic. However, the desired and achieved integration of several theoretical aspects from financial and economic theory is of course accompanied by an omittance of details in each of them. Nevertheless, the different layers of the general equilibrium model still accumulate to a considerably complex framework – not to mention the duplicity required to thoroughly integrate human labor and its capitalized value.

6.2 Outlook on Pensionomics

This contribution has addressed the portfolio perspective of the two financing mechanisms in the multi-pillar pension system and has focussed on the principle of efficiency in the management of the associated risks. By considering only a very stylized form, it has neglected major institutional aspects of actual pension systems as well as any transition problems. Instead of such incremental analysis for fixing real-world pay-as-you-go systems, the pension arrangements have been envisaged so that they would not need any fixing.

Thereby, the analysis is in line with the visionary consideration of the PAYGO mechanism as a mean of risk sharing.

Risk Sharing

The risk sharing understanding on unfunded pension mechanism is not new. Above all Robert J. Shiller emphasizes and advocates this role of the pay-as-you-go mechanism:[2] the central function of the pension system is managing the risk that some random events could leave old people impoverished. The notions of "old-age insurance" and "social security" highlight this risk-management function. Other functions of actual PAYGO systems can be handled by other governmental means, like programs to encourage individual saving, general welfare programs, or progressive income taxes. The presented analysis has agreed with this view and hence focussed exclusively on the risk management role.

Because management of risks does not make the underlying risks disappear, the risks must still be borne by someone. The benefit of insurance and risk management is that individual agents can escape risks that are concentrated in their hands and spread these risks over many agents so that the negative effects on the individual are marginalized. Therefore, risk management implies always risk sharing. The discussion in Sect. 5.4.1 illustrated this with the idiosyncratic risk components of each kind of security. Without tradable human capital or PAYGO, agents of the middle-aged and old generation are fully exposed to the factor share risk of human labor and physical capital respectively. Marketability, or alternatively the calibrated replicating PAYGO system, has allowed for a reduction of these risk concentrations. Due to the simplifying assumptions of the framework, like that of a closed economy, the factor share risk was not only marginalized but could even be eliminated entirely. Still, the endowment risk remains undiversifiable and must thus be borne by all generations. Being a national embodiment of ancient family traditions and obligations, this form of *intergenerational* risk sharing by PAYGO has naturally gained the most attention, and Otto von Bismarck's old age provision plan for Germany can be seen as its first real implementation.[3] As Shiller [2003*a*, p. 259-262] illustrates, the Bismarckian realization of the risk sharing argument was only possible with the strict German bureaucratical organization – at that time the most advanced form of public infrastructure.

[2]See Shiller [1998, p. 1] and Shiller [2003*a*].

[3]In addition to this, Shiller [1998] emphasizes two more dimensions of societal risk sharing. An *intragenerational* dimension of risk sharing is justified by similar reasons of market failure as the intergenerational one: when risks materialize before agents are old, or before they are capable of making risk management arrangements for themselves, only public risk-management devices can effectively allow elderly to pool the chance that some of them have higher income than others. And with open economies, there is the possibility of *international* risk sharing. Shiller [1993] gives the theoretical basis for such international trading in different components of national income.

But with the vast advances of modern information technology other means of societal risk management should be viable. In Shiller's [2003*a*] visionary view of a "New Financial Order" improved data availability will ultimately allow societies to trade other components of national income than that of physical capital and to allocate so the associated risks more effectively and efficiently. Shiller considers this as a democratization of finance, because the virtue of its tools and risk-management devices is not limited to a small group of wealthy individuals active in financial markets of the old-fashioned understanding anymore, but applicable to other forms of income and available for all members of society. Tradability of human labor would be such a democratization. With marketable human capital, the diversification solution derived in Chap. 3 would not be a hypothetical one anymore. Instead, the actual first-best allocation might become directly realizable, if *all* agents could indeed borrow against their future wage incomes because wide-spread data technology, innovative indexation and modern insurance contracts overcome the traditional moral hazard problems. Without markets for different classes of human capital failing, there will be no role for Chap. 4's replication with PAYGO – at least from the multi-pillar portfolio perspective of risk diversification. While the realization of this vision will still take considerable time, the traditional pay-as-you-go mechanisms could remain the single most important device for such societal risk management. As an already realized second-best, the old-fashioned PAYGO systems can help to enhance societies' handling of risks. However, in order to be considered as a trustworthy and sustainable tool of risk management by its participants, governmental bodies must overcome the traditional view of PAYGO as the sole arrangement for old-age and optimize its functioning with respect to the modern idea of risk sharing by substantial redesigning.

With risk management as a central function of pension systems, it would appear that the theoretical issues involved in the design of pay-as-you-go systems are fundamentally in the area of finance. However, the finance literature generally presumes that the underlying issue is the individual's problem of optimizing a portfolio of investment assets. But, in designing pension mechanisms, the fundamental problem is the social planer's or government's problem of deciding on institutions on behalf of the individuals. These institutions must thus promote the management of risks not only among people who are as rational investors actively managing their own risks, but also among people who are still minors or not even born yet. Furthermore, limits in the individuals' ability of financial management and, above all, limitations of the existing financial institutions and markets, require the theoretical approach for the designing of pension systems to be broader than the pure financial one. Since the holistic view of a social planer does not need to take as given the individual wealth positions or the investment and hedging vehicles that are available, the problem of old age insurance systems differs substantially from the typical problems of theoretical finance. Chapter 2 has reflected this view: the pension problem requires a perspective that is more comprehensive and fundamental than the standard problem of individual optimal portfolio management. It has been argued that the consequence should be a combination of theories from finance and economics and possibly other areas. In this sense, a step towards Pensionomics could be considered as a form of Shiller's [2003*a*] democratization of finance.[4]

Pensionomics

Of course, an *exhaustive* answer the question of how to effectively and efficiently deliver retirement income in face of global aging requires many further steps into Pensionomics. This work tried to establish some reference point based solely on efficiency and risk considerations. In this regard an important caveat is inevitable: even establishing Pensionomics with building blocks from past research might miss the future development similar to the thankfully unrealized predictions of Malthus [1798] or Meadows et al. [1972, 1974]. Even though the past development is now not extrapolated numerically, the conceptual procedure still rests heavily on the understanding of finance and economics built in an era of physical capital's scarcity. Yet, as Chap. 1 emphasized, from the standpoint of economic activity in the future, a different, but likely even more crucial, natural resource is becoming relatively scarce in course of the Demographic Transition: human beings. This shift might gradually produce absolutely unexpected relations in the functioning of an economy. The core of all the economic theory since Malthus [1798] can be summarized in a single sentence: the more people are using a stock of resources the lower the income per person, *if all else equal.* Exactly this "all else equal" seems to be the most important question for Pensionomics. For centuries, mankind managed to escape the law of diminishing returns by progress and technological advances and avoided so the Malthusian doom or the Club of Rome's projections. This possibility also exists in the risk sharing and diversification argument of the presented replication framework: if the total factor productivity of Chaps. 3 and 4 – or Sect. 5.3's level of technology – is sufficiently growing, this trend can overcompensate any negative shocks in the production elasticities of individual factors and a negative demographic development by a shrinking working population. In this sense, it does not only represent an endowment risk but also a substantial chance to mitigate the consequences of global aging. If overall national income grows due to some third influence factor, scarcity of the pure inputs of physical capital and human labor is less severe.

Consequently, the very interesting question for the pension issue and general future economic development is: how will aging affect technological progress? Knowledge in this area is very limited. The explanation of technological progress alone presents a difficult task on its own; assessing the impacts of

[4] See Shiller [1998, p. 1-2 and 10-16]. Besides Merton [1983], similar manifestations of the intergenerational risk sharing perspective on PAYGO include Bohn [1998], Demange and Laroque [1999], or Krueger and Kubler [2001].

aging on it adds additional complexity. Though the literature shows some empirical support for an effect of the demographic structure on economic growth,[5] the evidence for such an effect is not robust and even the direction of any effect remains uncertain: a more aged workforce can either be less dynamic and innovative and hence have a negative impact on productivity growth or, alternatively, the expected relative labor scarcity may act as a stimulus to technical advances. Both arguments seem equally intuitive and reasonable: on the one hand the number of young researchers and their ideas might be the crucial factor for progress, on the other hand, extended research and development typically only starts when a resource's limitation becomes binding, so that savings in it yield economic profits. Whichever of the hypothesis one chooses, the selection will be critical for future trends in technical progress and thus for long-term macroeconomic projections. Still, practical recommendations to politics seem clear: as for other economic problems, enhanced growth due to progress and technological advance will ease any adjustments and changes that are required to face the consequences of population aging. Research in this direction must be fostered. Furthermore, fast aging societies should overcome their attitudes towards the elderly in the working force: despite some particularities they present a valuable and hardly used resource for macroeconomic production. The old – as well as females – must be better integrated into the working force so that a higher effective utilization rate of the potential work force buffers the general trend of population decline.
Another shortcoming of this work and an important aspect for the future development are the interactions between economic variables and fertility. The exogenous population change inherent in the neoclassical models is clearly unsatisfactory, because it neglects them completely. But even the theory of the Demographic Transition, which has overcome this problem, lacks predictions beyond its forth stage of reduced mortality and fertility. Furthermore, there is the chance of purposeful policy changes to influence future demographic development, even though this would happen only gradually and can hardly revert the general trend. The New Home Economics argue that the parental "demand" for children has declined, since the private opportunity costs of children have risen with increasing opportunity costs of forgone family income or higher education costs. For an aging society, however, additional babies do have a high marginal benefit. Politics should internalize this externality to the family by alterations in the tax code and calculation of PAYGO benefits, focussed government spending on childcare, or other means dedicated to lower the private costs of children. Political progress, in this sense, would also change the "all else equal" and help to mitigate the consequences of the negative demographic trend.

[5] See, for instance, Simon and Zinsmeister [n.d.] or Kögel [2005] and references therein.

Deficit Delusion

The important, but in the end unsurprising result, of this contribution is that the portfolio perspective on the two pension financing mechanisms can only help to manage the risk associated with the generation of national income, but not eliminate the risk itself. From an aggregate perspective, even a properly designed public pension system can only have an insurance character with respect to the factor share risk, but not with respect to the general level of national income. Hence, the risk sharing argument can only refer to the national income available for all generations alive at the same time. From the viewpoint of portfolio optimization, benefits must be defined in such a way that they refer to this available income and do not extrapolate past shares of some previous national income. Savings in form of financial assets have no value per se, if the working generation is not demanding the service of the underlying real capital and is not willing to pay for it. In this sense, the form of pension financing – funded or unfunded – does not matter. Retirees do actually not live off their previous savings, but off the current production of consumable output which the working generation is producing under usage of the accumulated real capital.
As in the Tweedles' case, the mere naming of these intergenerational relations does not matter. Only a thorough intertemporal analysis allows to assess the net effects on different generations. With taxation of the young and transfers to the old, Tweedledum's country is effectively operating a pay-as-you-go pension system. Tweedledee's country, on the other hand, allows its citizens to accumulate some funded savings to live off as elderly. The table in the introduction illustrated that the numerics are the same. At all times, generations of all ages must share national income – or gross domestic product – of that period. The soundness of the public sector, of the fiscal policy and of other collective schemes like the pay-as-you-go mechanism, can only be assessed from a holistic point of view. Single replacement rate extrapolations or debt-to-income ratios, which ignore economic reality and future limitations, are useless.

> *Alice in Fiscal Land*
>
> *As far as we know, poor Alice is still sitting in a dungeon explaining to the Queen's guards that in and of itself, the level of government debt and its change over time tell us nothing whatsoever about a country's fiscal policy. For people, particulary economists, who have spent their entire adult lives discussing fiscal policy in terms of the size of the government's deficit and comparing countries' debt-to-GDP ratios, this proposition, if true, should be highly disturbing. It's true.*
>
> *Kotlikoff and Burns [2004]: "The Coming Generational Storm", p. 76*

Part IV

Appendices

A

Methodological Foundation

A.1 Asset Pricing

Based on the intertemporal optimization underlying the Fisher-Separation as illustrated in Fig. 2.2, the stochastic discount factor can be understood as an integration of this rationale across different states of nature.
Consider Fisher's [1930] approach, which is a problem *without* uncertainty, for an investor maximizing utility U of consumption $c(0), c(1)$ in a simplified two-period version of the problem of (2.2.1) and (2.2.2):

$$\max U = u\left(c(0)\right) + \beta \times u\left(c(1)\right) \tag{A.1.1}$$

subject to his normalized budget condition

$$1 = c(0) + \frac{c(1)}{R(1)}. \tag{A.1.2}$$

Limiting the analysis on two cases, the gross return on the capital market, $R(1)$, shall be characterized by the possibility of an up- and a down-state with the probabilities P_u and $P_d = 1 - P_u$. In each of them, the investor optimally trades off future and present consumption, so that his state dependent first order conditions are given by

$$\frac{\partial U}{\partial c_u(0)} = \beta \frac{\partial U}{\partial c_u(1)} \times R_u(1) \quad \text{in the up-state “u”,} \tag{A.1.3}$$

$$\frac{\partial U}{\partial c_d(0)} = \beta \frac{\partial U}{\partial c_d(1)} \times R_d(1) \quad \text{in the down-state “d”.} \tag{A.1.4}$$

One can multiply (A.1.3) with P_u as well as (A.1.4) with $P_d = 1 - P_u$ and sum the two results on the left and right hand side to

$$\frac{\partial U}{\partial c_u(0)} P_u + \frac{\partial U}{\partial c_d(0)} \left(1 - P_u\right) = \beta \frac{\partial U}{\partial c_u(1)} R_u(1) P_u + \beta \frac{\partial U}{\partial c_d(1)} R_d(1) \left(1 - P_u\right) \tag{A.1.5}$$

Noting now that the investor must decide on his initial consumption, $c(0)$, *independent* of the later return realization, it must hold $c_u(0) = c_d(0) = c(0)$ and hence $\partial U/\partial c_u(0) = \partial U/\partial c_d(0) = \partial U/\partial c(0)$. Using this and applying the expectation operator $\mathbb{E}$ on the right hand side of (A.1.5) yields

$$\frac{\partial U}{\partial c(0)} = \mathbb{E}\left[\beta \frac{\partial U}{\partial \tilde{c}(1)} \tilde{R}(1)\right].$$

Dividing by $\partial U/\partial c(0)$ gives

$$1 = \mathbb{E}\left[\beta \frac{\frac{\partial U}{\partial \tilde{c}(1)}}{\frac{\partial U}{\partial c(0)}} \tilde{R}(1)\right], \tag{A.1.6}$$

which is the central pricing equation for the uncertain return $\tilde{R}(1)$ analogous to (2.2.7), only missing the corresponding definition of the SDF.

A.2 Growth Theory

Based on the relations in absolute terms from Sect. 2.2.5 the derivation of the per-capita expression is straightforward. Since (2.2.9) also holds for $\lambda = 1/L(t)$ output can be written as

$$Y(t) = F\left(K(t), L(t)\right) = L(t) \times F\left(\frac{K(t)}{L(t)}, 1\right) = L(t) f(k), \tag{A.2.1}$$

where $k(t) \equiv K(t)/L(t)$ is real capital per worker, and the function $f\left(k(t)\right)$ is defined to equal $F(k(t), 1)$. With additionally $y(t) \equiv Y(t)/L(t)$ as output per worker one can hence write the production function in its *intensive form*

$$y(t) = f\left(k(t)\right). \tag{A.2.2}$$

These per-capita relations allow then to derive the fundamental equation of the Solow-Swan-Model: dividing both sides of (2.2.13) by $L(t)$ first, and replacing then $(dK/dt)/L(t)$ with $dk(t)/dt + nk$ from the derivative of $k(t)$ with respect to time, one gets

$$\frac{dk(t)}{dt} = s \times f\left(k(t)\right) - n \times k(t). \tag{A.2.3}$$

This is (2.2.14) in the text.

B

Tradable Human Capital

B.1 Optimization by Agents

B.1.1 Optimization by Middle-aged

By the mechanics of stochastic dynamic programming one can substitute the results of (3.4.14) for $U_o(t+1)$ in (3.4.15) to obtain

$$\begin{aligned} J_m[w_m(t),t] &= \max_{\substack{c_m(t),\\ \alpha_{K,m}(t),\\ \alpha_{H,m}(t)}} \left\{ \ln c_m(t) + \mathbb{E}_t \left[\beta U_o[t+1] \right] \right\} \\ &= \max_{\bullet} \left\{ \ln c_m(t) + \mathbb{E}_t \left[\beta J_o[w_o(t+1), t+1] \right] \right\} \\ &= \max_{\bullet} \left\{ \ln c_m(t) + \beta \mathbb{E}_t \left[\ln w_o(t+1) \right] \right\} \end{aligned} \tag{B.1.1}$$

Using now the wealth definition (3.4.6) and the budget constraint (3.4.10) for $w_o(t+1)$ gives (3.4.16):

$$J_m[w_m(t),t] =$$

$$\max_{\substack{c_m(t)\\ \alpha_{K,m}(t)\\ \alpha_{H,m}(t)}} \left\{ \begin{array}{l} \ln c_m(t) + \\ \beta \mathbb{E}_t \left[\ln \left(\begin{bmatrix} w_m(t) \\ - c_m(t) \end{bmatrix} \left\{ \begin{array}{c} \alpha_{K,m}(t)[\tilde{R}_K(t+1) - R(t)] \\ + \alpha_{H,m}(t)[\tilde{R}_H(t+1) - R(t)] \\ + R(t) \end{array} \right\} \right) \right] \end{array} \right\} . \tag{B.1.2}$$

This is (3.4.16) in the text.
In order to show (3.4.23), take the budget constraint (3.4.10) and apply the savings policy (3.4.20) and the definition of $\tilde{R}^*_{P,m}(t+1)$. At the optimum, the middle-aged agent transfer consumption opportunities characterized by

$$w_o^*(t+1) = s_m^*(t) \left\{ \begin{array}{l} \alpha_{K,m}^*(t) \left[\tilde{R}_K(t+1) - R(t) \right] \\ + \alpha_{H,m}^*(t) \left[\tilde{R}_H(t+1) - R(t) \right] \\ + R(t) \end{array} \right\}$$
$$= \frac{\beta}{\beta+1} w_m(t)\ \tilde{R}_{P,m}^*(t+1). \qquad \text{(B.1.3)}$$

Using optimal values allows to drop the maximization in (B.1.1). Applying (B.1.3) as well as (3.4.19) and simplifying gives then

$$J_m[w_m(t),t] = \ln\left(\frac{1}{\beta+1} w_m(t)\right) + \beta \mathbb{E}_t \left[\ln\left(\frac{\beta}{\beta+1} w_m(t) \tilde{R}_{P,m}^*(t+1)\right)\right]$$
$$= (\beta+1)\ln w_m(t) + \beta \ln \beta - (\beta+1)\ln(\beta+1) + \beta \mathbb{E}_t \left[\ln \tilde{R}_{P,m}^*(t+1) \right]. \qquad \text{(B.1.4)}$$

This is (3.4.23) in the main text.

B.1.2 Optimization by Young

The derivation of (3.4.25) is analogous to that of (3.4.16). Start with applying the principle of stochastic dynamic programming in (3.4.24)

$$J_y[w_y(t),t] = \max_{\substack{c_y(t), \\ \alpha_{K,y}(t), \\ \alpha_{H,y}(t)}} \left\{ \ln c_y(t) + \mathbb{E}_t \left[\beta U_m(t+1) \right] \right\}$$
$$= \max_{\bullet} \left\{ \ln c_y(t) + \beta \mathbb{E}_t \left[J_m[w_m(t+1), t+1] \right] \right\}$$

and replace $J_m[w_m(t+1), t+1]$ with the results of (3.4.23)

$$= \max_{\bullet} \left\{ \ln c_y(t) + \beta \mathbb{E}_t \left[\begin{array}{l} \beta \ln \beta - (\beta+1)\ln(\beta+1) \\ + (\beta+1)\ln w_m(t+1) \\ + \beta \mathbb{E}_{t+1} \left[\ln \tilde{R}_{T,m}^*(t+2) \right] \end{array} \right] \right\}.$$

Applying the law of iterative expectations allows to simplify this

$$= \max_{\bullet} \left\{ \begin{array}{l} \ln c_y(t) - \beta(\beta+1)\ln(\beta+1) + \beta^2 \ln \beta \\ + \beta^2 \mathbb{E}_t \left[\ln \tilde{R}_{P,m}^*(t+2) \right] \\ + \beta(\beta+1)\mathbb{E}_t \left[\ln w_m(t+1) \right] \end{array} \right\}. \qquad \text{(B.1.5)}$$

After substituting (3.4.9) for $w_m(t+1)$ into (B.1.5) one obtains the main text's (3.4.25):

$$J_y[w_y(t),t] = \max_{\substack{c_y(t)\\ \alpha_{K,y}(t)\\ \alpha_{H,y}(t)}} \left\{ \begin{array}{l} \ln c_y(t) - \beta(\beta+1)\ln(\beta+1) + \beta^2 \ln\beta \\ + \beta^2 \mathbb{E}_t \left[\ln \tilde{R}^*_{P,m}(t+2)\right] + \beta(\beta+1)\times \\ \mathbb{E}_t \left[\ln \left(\begin{bmatrix} w_y(t) \\ - c_y(t) \end{bmatrix} \times \left\{ \begin{array}{l} \alpha_{K,y}(t)[\tilde{R}_K(t+1) - R(t)] \\ + \alpha_{H,y}(t)[\tilde{R}_H(t+1) - R(t)] \\ + R(t) \end{array} \right\} \right) \right] \end{array} \right\}. \tag{B.1.6}$$

For (3.4.32), start with the above equation and plug in the optimal $c_y^*(t)$ from (3.4.28), $s_y^*(t)$ from (3.4.29) as well as $\tilde{R}^*_{P,y}(t+1)$ from (3.4.30). Usage of these optimal policies makes the maximization redundant:

$$J_y[w_y(t),t] = \left\{ \begin{array}{l} \ln \left(\frac{1}{\beta^2+\beta+1} w_y(t) \right) - \beta(\beta+1)\ln(\beta+1) + \beta^2 \ln\beta \\ + \beta^2 \mathbb{E}_t \left[\ln \tilde{R}^*_{P,m}(t+2)\right] \\ + \beta(\beta+1)\mathbb{E}_t \left[\ln \left(\frac{\beta^2+\beta}{\beta^2+\beta+1} w_y(t) \tilde{R}^*_{P,y}(t+1) \right) \right] \end{array} \right\}. \tag{B.1.7}$$

Simplifying the terms in $\{\circ\}$

$$\begin{aligned} &\ln \left(\frac{1}{\beta^2+\beta+1} w_y(t) \right) - \beta(\beta+1)\ln(\beta+1) + \beta^2 \ln\beta \\ &\quad + \beta^2 \mathbb{E}_t \left[\ln \tilde{R}^*_{P,m}(t+2)\right] + \beta(\beta+1)\mathbb{E}_t \left[\ln \left(\frac{\beta^2+\beta}{\beta^2+\beta+1} w_y(t) \tilde{R}^*_{P,y}(t+1) \right) \right] \\ &= -\ln(\beta^2+\beta+1) - \beta(\beta+1)\ln(\beta+1) + \beta^2\ln\beta + \ln w_y(t) \\ &\quad + \beta^2 \mathbb{E}_t \left[\ln \tilde{R}^*_{P,m}(t+2)\right] + \beta(\beta+1)\ln\beta\ln(\beta+1) - (\beta^2+\beta)\ln(\beta^2+\beta+1) \\ &\quad + \beta(\beta+1)\ln w_y(t) + \beta(\beta+1)\mathbb{E}_t \left[\ln \tilde{R}^*_{P,y}(t+1)\right] \end{aligned}$$

and rearranging finally yields to (3.4.32) from the text

$$J_y[w_y(t),t] = \left\{ \begin{array}{l} \beta^2 \ln\beta + \beta(\beta+1)(\ln\beta - 1)\ln(\beta+1) \\ - (\beta^2+\beta+1)\ln(\beta^2+\beta+1) + (\beta^2+\beta+1)\ln w_y(t) \\ + \beta(\beta+1)\mathbb{E}_t \left[\ln \tilde{R}^*_{P,y}(t+1)\right] + \beta^2 \mathbb{E}_t \left[\ln \tilde{R}^*_{P,m}(t+2)\right] \end{array} \right\}. \tag{B.1.8}$$

B.1.3 Stochastic Discount Factor

SDF of Middle-Aged

The derivation of (3.4.36) is analogous to the one of (3.4.35). Start with a middle-aged agent's SDF from (3.4.34), use (3.4.12) to replace $\tilde{c}_o^*(t+1)$, plug in (3.4.19) for $c_m^*(t)$ as well as (3.4.10) for $\tilde{w}_o(t+1)$:

$$\tilde{M}_m(t+1) = \beta \frac{c_m^*(t)}{\tilde{c}_o^*(t+1)} = \beta \frac{\frac{1}{\beta+1} w_m(t)}{\tilde{w}_o(t+1)} = \beta \frac{\frac{1}{\beta+1} w_m(t)}{s_m^*(t)\tilde{R}_{P,m}^*(t+1)},$$

due to optimality. Finally, substitute (3.4.20) for $s_m^*(t)$ and simplify

$$= \beta \frac{\frac{1}{\beta+1} w_m(t)}{\frac{\beta}{\beta+1} w_m(t)\tilde{R}_{P,m}^*(t+1)} = \frac{1}{\tilde{R}_{P,m}^*(t+1)}. \tag{B.1.9}$$

This corresponds to (3.4.36) in the text.

Alternative Derivation of SDFs

An alternative way to establish the stochastic discount factors is the application of the envelope condition from the Bellmann principle. The envelope condition says that the derivative of the indirect utility function with respect to the indirect state variable equals the derivative of the period utility function with respect to the control, i.e.

$$\frac{\partial J_a}{\partial w_a} = \frac{\partial u_a}{\partial c_a}. \tag{B.1.10}$$

Using (B.1.10) in the definition of the stochastic discount factor allows to write this in case of a young agent as

$$\tilde{M}_y(t+1) \equiv \frac{\frac{\partial U_y}{\partial \tilde{c}_m(t+1)}}{\frac{\partial U_y}{\partial c_y(t)}} = \beta \frac{\frac{\partial u_m}{\partial \tilde{c}_m(t+1)}}{\frac{\partial u_y}{\partial c_y(t)}} = \beta \frac{\frac{\partial J_m}{\partial w_m(t+1)}}{\frac{\partial J_y}{\partial w_y(t)}}, \tag{B.1.11}$$

and for a middle-aged agent as

$$\tilde{M}_m(t+1) \equiv \frac{\frac{\partial U_m}{\partial \tilde{c}_o(t+1)}}{\frac{\partial U_m}{\partial c_m(t)}} = \beta \frac{\frac{\partial u_o}{\partial \tilde{c}_o(t+1)}}{\frac{\partial u_m}{\partial c_m(t)}} = \beta \frac{\frac{\partial J_o}{\partial w_o(t+1)}}{\frac{\partial J_m}{\partial w_m(t)}}. \tag{B.1.12}$$

According to (3.4.14), (3.4.23) and (3.4.32) the required derivatives are

$$\frac{\partial J_y}{\partial w_y(t)} = \frac{\beta^2+\beta+1}{w_y(t)}, \quad \frac{\partial J_m}{\partial w_m(t)} = \frac{\beta+1}{w_m(t)}, \quad \text{and} \quad \frac{\partial J_o}{\partial w_o(t)} = \frac{1}{w_o(t)}. \tag{B.1.13}$$

Applying (B.1.13) in (B.1.11), respectively (B.1.12), yields

$$\tilde{M}_y(t+1) = \beta \frac{\frac{\partial J_m}{\partial w_m(t+1)}}{\frac{\partial J_y}{\partial w_y(t)}} = \beta \frac{\frac{\beta+1}{w_m(t+1)}}{\frac{\beta^2+\beta+1}{w_y(t)}} = \beta \frac{\beta+1}{\beta^2+\beta+1} \frac{w_y(t)}{w_m(t+1)} \quad \text{(B.1.14)}$$

$$\tilde{M}_m(t+1) = \beta \frac{\frac{\partial J_o}{\partial w_o(t+1)}}{\frac{\partial J_m}{\partial w_m(t)}} = \beta \frac{\frac{1}{w_o(t+1)}}{\frac{\beta+1}{w_m(t)}} = \beta \frac{1}{\beta+1} \frac{w_m(t)}{w_o(t+1)} . \quad \text{(B.1.15)}$$

Replacing now $w_a(t+1)$ with $s_a^*(t) \times \tilde{R}_{P,a}^*(t+1)$ in each formulation and then substituting (3.4.20), respectively (3.4.29), for the savings gives:

$$\begin{aligned} \tilde{M}_y(t+1) &= \beta \frac{\beta+1}{\beta^2+\beta+1} \frac{w_y(t)}{w_m(t+1)} = \beta \frac{\beta+1}{\beta^2+\beta+1} \frac{w_y(t)}{s_y^*(t)\tilde{R}_{P,y}^*(t+1)} \\ &= \beta \frac{\beta+1}{\beta^2+\beta+1} \frac{w_y(t)}{\frac{\beta^2+\beta}{\beta^2+\beta+1} w_y(t)\tilde{R}_{P,y}^*(t+1)} = \frac{1}{\tilde{R}_{P,y}^*(t+1)} \end{aligned} \quad \text{(B.1.16)}$$

$$\begin{aligned} \tilde{M}_m(t+1) &= \beta \frac{1}{\beta+1} \frac{w_m(t)}{w_o(t+1)} = \beta \frac{1}{\beta+1} \frac{w_m(t)}{s_m^*(t)\tilde{R}_{P,m}^*(t+1)} \\ &= \beta \frac{1}{\beta+1} \frac{w_m(t)}{\frac{\beta}{\beta+1} w_m(t)\tilde{R}_{P,m}^*(t+1)} = \frac{1}{\tilde{R}_{P,m}^*(t+1)} . \end{aligned} \quad \text{(B.1.17)}$$

These are – of course – the same stochastic discount factors as (3.4.35) and (3.4.36) in the text.

B.2 Capital Market Equilibrium

B.2.1 Deriving Equilibrium

Derivation of (3.5.12)

Inserting (3.5.8) and (3.5.9) in the portfolio rules, (3.4.18) or (3.4.27), yields for each security $i = H, K$

$$0 = \mathbb{E}_t \left[\frac{\tilde{R}_i(t+1) - R(t)}{\alpha_K^M(t)[\tilde{R}_K(t+1) - R(t)] + \alpha_H^M(t)[\tilde{R}_H(t+1) - R(t)] + R(t)} \right] .$$

Using then (3.5.10) and (3.5.11) gives after some manipulations

$$\begin{aligned} &= \mathbb{E}_t \left[\frac{\tilde{R}_i(t+1) - R(t)}{\frac{I(t)[\tilde{R}_K(t+1)-R(t)]}{I(t)+\vartheta N_y(t)p_H(t)} + \frac{\vartheta N_y(t)p_H(t)[\tilde{R}_H(t+1)-R(t)]}{I(t)+\vartheta N_y(t)p_H(t)} + \frac{I(t)+\vartheta N_y(t)p_H(t)R(t)}{I(t)+\vartheta N_y(t)p_H(t)}} \right] \\ &= \mathbb{E}_t \left[\frac{[I(t) + \vartheta N_y(t)p_H(t)][\tilde{R}_i(t+1) - R(t)]}{I(t)\tilde{R}_K(t+1) + \vartheta N_y(t)p_H(t)\tilde{R}_H(t+1)} \right] . \end{aligned}$$

Involving now the return definitions, (3.2.14) and (3.3.9),

$$= \mathbb{E}_t \left[\frac{[I(t) + \vartheta N_y(t) p_H(t)][\tilde{R}_i(t+1) - R(t)]}{I(t)\frac{\vartheta\tilde{\theta}(t+1)\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}{I(t)} + \vartheta N_y(t) p_H(t) \frac{\vartheta(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{\vartheta N_y(t) p_H(t)}} \right]$$

$$= \ldots = \mathbb{E}_t \left[\frac{\tilde{R}_i(t+1) - R(t)}{\frac{\vartheta\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}{I(t)+\vartheta N_y(t) p_H(t)}} \right]$$

and applying (3.2.12) as well as (3.3.10) for the return on the market, one finally obtains

$$0 = \mathbb{E}_t \left[\frac{\tilde{R}_i(t+1) - R(t)}{\tilde{R}_M(t+1)} \right]. \tag{B.2.1}$$

This proofs (3.5.12) in the text.

Derivation of (3.5.14)

Rearranging (B.2.1) by splitting the fraction in two summands, separating the expectations and taking one term on the other side yields

$$\mathbb{E}_t \left[\frac{\tilde{R}_K(t+1)}{\tilde{R}_M(t+1)} \right] = \mathbb{E}_t \left[\frac{\tilde{R}_H(t+1)}{\tilde{R}_M(t+1)} \right] = \mathbb{E}_t \left[\frac{R(t)}{\tilde{R}_M(t+1)} \right]. \tag{B.2.2}$$

Then, by multiplication with the corresponding market portfolio weight one has

$$\mathbb{E}_t \left[\frac{\alpha_K^M(t)\tilde{R}_K(t+1)}{\tilde{R}_M(t+1)} \right] = \mathbb{E}_t \left[\frac{\alpha_K^M(t) R(t)}{\tilde{R}_M(t+1)} \right] \tag{B.2.3}$$

and
$$\mathbb{E}_t \left[\frac{\alpha_H^M(t)\tilde{R}_H(t+1)}{\tilde{R}_M(t+1)} \right] = \mathbb{E}_t \left[\frac{\alpha_H^M(t) R(t)}{\tilde{R}_M(t+1)} \right]. \tag{B.2.4}$$

Adding (B.2.3) and (B.2.4) on each side gives with $\alpha_H^M(t) = 1 - \alpha_K^M(t)$

$$\mathbb{E}_t \left[\frac{\alpha_K^M(t)\tilde{R}_K(t+1) + \alpha_H^M(t)\tilde{R}_H(t+1)}{\tilde{R}_M(t+1)} \right] = \mathbb{E}_t \left[\frac{(\alpha_K^M(t) + \alpha_H^M(t))R(t)}{\tilde{R}_M(t+1)} \right]$$

$$\mathbb{E}_t \left[\frac{\tilde{R}_M(t+1)}{\tilde{R}_M(t+1)} \right] = \mathbb{E}_t \left[\frac{R(t)}{\tilde{R}_M(t+1)} \right]. \tag{B.2.5}$$

By canceling the fraction and combing (B.2.2) with (B.2.5) one obtains (3.5.14) from the text:

$$1 = \mathbb{E}_t \left[\frac{R(t)}{\tilde{R}_M(t+1)} \right] = \mathbb{E}_t \left[\frac{\tilde{R}_K(t+1)}{\tilde{R}_M(t+1)} \right] = \mathbb{E}_t \left[\frac{\tilde{R}_H(t+1)}{\tilde{R}_M(t+1)} \right]. \tag{B.2.6}$$

B.2.2 Market Portfolio Weights

Derivation of (3.5.15)

To derive (3.5.15), start with (3.5.14), insert the derivations for the return on the market – from (3.3.10) – and for the human capital security – from (3.3.9) – and then simplify:

$$
\begin{aligned}
1 &= \mathbb{E}_t\left[\frac{\tilde{R}_H(t+1)}{\tilde{R}_M(t+1)}\right] = \mathbb{E}_t\left[\frac{\frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{N_y(t)p_H(t)}}{\frac{\vartheta\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}{I(t)+\vartheta N_y(t)p_H(t)}}\right] \\
&= \mathbb{E}_t\left[\frac{\vartheta(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{\vartheta N_y(t)p_H(t)}\frac{I(t)+\vartheta N_y(t)p_H(t)}{\vartheta\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}\right] \\
&= \mathbb{E}_t\left[\frac{\vartheta(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{\vartheta\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}\frac{I(t)+\vartheta N_y(t)p_H(t)}{\vartheta N_y(t)p_H(t)}\right] \\
&= \mathbb{E}_t\left[\frac{(1-\tilde{\theta}(t+1))\tilde{Y}(t+1)}{\tilde{Y}(t+1)+\frac{1-\vartheta}{\vartheta}\tilde{I}(t+1)}\right]\left[1+\frac{I(t)}{\vartheta N_y(t)p_H(t)}\right].
\end{aligned}
\tag{B.2.7}
$$

Relation of $\mathbb{E}\left[\tilde{\psi}\right]$ *vs.* $\bar{\psi}$

The relation of $\bar{\psi}$ and the true mean of $\tilde{\psi}$ is influenced by Jensen's inequality. As $(1+\frac{1-\vartheta}{\vartheta}\tilde{\psi})^{-1}$ is a convex function of $\tilde{\psi}$ it holds

$$\mathbb{E}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}}\right] > \frac{1}{1+\frac{1-\vartheta}{\vartheta}\mathbb{E}\left[\tilde{\psi}\right]}.$$

Now let the inequality be corrected by a factor $\lambda > 1$ such that

$$\mathbb{E}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}}\right] = \frac{\lambda}{1+\frac{1-\vartheta}{\vartheta}\mathbb{E}\left[\tilde{\psi}\right]},$$

then $\bar{\psi}$ and the true mean, $\mathbb{E}\left[\tilde{\psi}\right]$, are related by

$$\bar{\psi} = \frac{\frac{1-\vartheta}{\vartheta}+\mathbb{E}\left[\tilde{\psi}\right]}{\lambda} - \frac{1-\vartheta}{\vartheta}. \tag{B.2.8}$$

Capitalization of Human Capital

Using the investment to output ratio, $\psi(t)$, and the covariance decomposition allows to rewrite (B.2.7) as

$$
\begin{aligned}
1 &= \mathbb{E}_t\left[\frac{1-\tilde{\theta}(t+1)}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}(t+1)}\right]\left[1+\frac{I(t)}{\vartheta N_y(t)p_H(t)}\right] \\
&= \left\{\mathbb{E}_t\left[1-\tilde{\theta}(t+1)\right]\mathbb{E}_t\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}(t+1)}\right]+\mathrm{Cov}\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}},\tilde{\theta}\right]\right\} \\
&\quad\times\left[1+\frac{I(t)}{\vartheta N_y(t)p_H(t)}\right].
\end{aligned}
$$

The assumptions of Sect. 3.2.1 and (3.5.16) allow now to reduce this to

$$
= \left[1-\bar{\theta}\right]\left[\frac{1}{1+\frac{1-\vartheta}{\vartheta}\bar{\psi}}\right]\left[1+\frac{I(t)}{\vartheta N_y(t)p_H(t)}\right]. \tag{B.2.9}
$$

Taking the second bracket to the left hand side and some further manipulations of (B.2.9), namely

$$
\begin{aligned}
1+\frac{1-\vartheta}{\vartheta}\bar{\psi} &= \left[1-\bar{\theta}\right]\left[1+\frac{I(t)}{\vartheta N_y(t)p_H(t)}\right] \\
\frac{1}{1-\bar{\theta}}+\frac{1-\vartheta}{\vartheta(1-\bar{\theta})}\bar{\psi}-1 &= \frac{I(t)}{\vartheta N_y(t)p_H(t)} \\
\frac{\vartheta+(1-\vartheta)\bar{\psi}-\vartheta(1-\bar{\theta})}{1-\bar{\theta}} &= \frac{I(t)}{N_y(t)p_H(t)},
\end{aligned}
$$

yield finally (3.5.17) of the text:

$$
N_y(t)p_H(t) = \frac{1-\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}\times I(t). \tag{B.2.10}
$$

Total Market Capitalization

To get the total market value, insert (3.5.18) in the definition of the total market capitalization from Sect. 3.3.4 and simplify

$$
\begin{aligned}
&K(t+1)p_K(t)+N_m(t+1)p_H(t) \\
&= K(t+1)p_K(t)+\frac{(1-\bar{\theta})\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}K(t+1)p_K(t) \\
&= \frac{(1-\bar{\theta})\vartheta+(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}K(t+1)p_K(t) = \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}K(t+1)p_K(t) \\
&= \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}\frac{1}{\vartheta}I(t).
\end{aligned}
$$

This is (3.5.19) in the main text.

Using Real Capital Return

One can also derive (3.5.17) using not the human capital return, but the real capital one. Yet, this is slightly more laborious. Starting again with (3.5.14), and inserting (3.2.14) gives:

$$
\begin{aligned}
1 &= \mathbb{E}_t\left[\frac{\tilde{R}_K(t+1)}{\tilde{R}_M(t+1)}\right] = \mathbb{E}_t\left[\frac{\frac{\vartheta\tilde{\theta}(t+1)\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}{I(t)}}{\frac{\vartheta\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}{I(t)+\vartheta N_y(t)p_H(t)}}\right] \\
&= \mathbb{E}_t\left[\frac{\vartheta\tilde{\theta}(t+1)\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}{I(t)}\frac{I(t)+\vartheta N_y(t)p_H(t)}{\vartheta\tilde{Y}(t+1)+(1-\vartheta)\tilde{I}(t+1)}\right] \\
&= \mathbb{E}_t\left[\frac{\tilde{\theta}(t+1)-1+1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}(t+1)}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}(t+1)}\right]\left[1+\frac{\vartheta N_y(t)p_H(t)}{I(t)}\right] \\
&= \mathbb{E}_t\left[\frac{\tilde{\theta}(t+1)-1}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}(t+1)}\right]\left[1+\frac{\vartheta N_y(t)p_H(t)}{I(t)}\right] \\
&\quad +\mathbb{E}_t\left[\frac{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}(t+1)}{1+\frac{1-\vartheta}{\vartheta}\tilde{\psi}(t+1)}\right]\left[1+\frac{\vartheta N_y(t)p_H(t)}{I(t)}\right] \\
&= [\bar{\theta}-1]\frac{1}{1+\frac{1-\vartheta}{\vartheta}\bar{\psi}}\frac{I(t)+\vartheta N_y(t)p_H(t)}{I(t)}+\frac{I(t)+\vartheta N_y(t)p_H(t)}{I(t)},
\end{aligned}
$$

where the last line follows due to (3.5.16). Multiplication of both sides with $I(t)/(I(t)+\vartheta N_y(t)p_H(t))$ yields

$$
\begin{aligned}
\frac{I(t)}{I(t)+\vartheta N_y(t)p_H(t)} &= \frac{\bar{\theta}-1}{1+\frac{1-\vartheta}{\vartheta}\bar{\psi}}+1 = \frac{\bar{\theta}+\frac{1-\vartheta}{\vartheta}\bar{\psi}}{1+\frac{1-\vartheta}{\vartheta}\bar{\psi}} \\
I(t) &= \frac{\vartheta\bar{\theta}+(1-\vartheta)\bar{\psi}}{\vartheta+(1-\vartheta)\bar{\psi}}\left[I(t)+\vartheta N_y(t)p_H(t)\right] \\
\frac{\vartheta+(1-\vartheta)\bar{\psi}}{\vartheta\bar{\theta}+(1-\vartheta)\bar{\psi}}I(t)-\frac{\vartheta\bar{\theta}+(1-\vartheta)\bar{\psi}}{\vartheta\bar{\theta}+(1-\vartheta)\bar{\psi}}I(t) &= \vartheta N_y(t)p_H(t) \\
\frac{1-\bar{\theta}}{\vartheta\bar{\theta}+(1-\vartheta)\bar{\psi}}I(t) &= N_y(t)p_H(t)
\end{aligned}
$$

which is the same result (3.5.17) as by usage of the human capital return.

B.3 General Equilibrium

B.3.1 Consumption

Simplification of (3.6.4) to (3.6.5) is straightforward. Starting with the original, extracting common terms, and summing them yields

$$S(t) = \frac{(1-\vartheta)\bar{\psi}+\vartheta}{(1-\vartheta)\vartheta\bar{\psi}+\vartheta^2\bar{\theta}} I(t) - \frac{1-\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} I(t) + \frac{1}{\beta^2+\beta+1}\frac{1-\bar{\theta}}{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}} I(t)$$

After simplifying and canceling terms one arrives at

$$S_m(t) = \left[\frac{1}{\vartheta} + \frac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)}\right] I(t), \tag{B.3.1}$$

which is (3.6.5) in the main text.

B.3.2 Investment-Output Ratio

Equation (3.6.10) is central to Chap. 3. To obtain it, start with (3.6.9) from the text:

$$1 - \left[\frac{(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}}{\vartheta} + \frac{1-\bar{\theta}}{\beta^2+\beta+1}\right]\frac{(1-\vartheta)\psi(t)+\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta}$$
$$= \left[\begin{array}{l} \dfrac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)} + \dfrac{1}{\beta\vartheta} \\ + \dfrac{1-\bar{\theta}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)\beta} + 1 \end{array}\right]\psi(t).$$

Summation of the fractions in brackets, splitting the last term in the first line, and taking all terms of $\psi(t)$ on the right hand side yields

$$1 - \frac{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)+(1-\bar{\theta})\vartheta}{(\beta^2+\beta+1)\vartheta}\frac{\vartheta}{(1-\vartheta)\bar{\psi}+\vartheta}$$
$$= \frac{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)+(1-\bar{\theta})\vartheta}{(\beta^2+\beta+1)\vartheta}\frac{(1-\vartheta)}{(1-\vartheta)\bar{\psi}+\vartheta}\psi(t)$$
$$+\frac{\left\{\begin{array}{l}(1-\bar{\theta})(\beta+1)\vartheta + \left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1) \\ +\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)\beta\vartheta\end{array}\right\}}{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)\beta\vartheta}\psi(t).$$

By multiplication of both sides with $\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right]\left[(1-\vartheta)\bar{\psi}+\vartheta\right](\beta^2+\beta+1)\beta\vartheta$ one gets

$$\left[\begin{array}{l}\left\{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right]\left[(1-\vartheta)\bar{\psi}+\vartheta\right](\beta^2+\beta+1)\beta\vartheta\right\} \\ -\left\{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)+(1-\bar{\theta})\vartheta\right\}\vartheta\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right]\beta\end{array}\right] = \psi(t)$$
$$\times\left[\begin{array}{l}\left\{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)+(1-\bar{\theta})\vartheta\right\}(1-\vartheta)\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right]\beta \\ +\left\{(1-\bar{\theta})(\beta+1)\vartheta\right\}\left[(1-\vartheta)\bar{\psi}+\vartheta\right] \\ +\left\{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)\right\}\left[(1-\vartheta)\bar{\psi}+\vartheta\right] \\ +\left\{\left[(1-\vartheta)\bar{\psi}+\vartheta\bar{\theta}\right](\beta^2+\beta+1)\beta\vartheta\right\}\left[(1-\vartheta)\bar{\psi}+\vartheta\right]\end{array}\right] \tag{B.3.2}$$

For simplicity the brackets on the left, [l.h.s.], and right hand side, [r.h.s.], of this equation will be simplified separately. Therefore write (B.3.2) as

$$[\text{l.h.s.}] = \psi(t) \times [\text{r.h.s.}] \tag{B.3.3}$$

The left hand side can be simplified by extracting common factors and simplifying

$$\begin{aligned}[\text{l.h.s.}] &= \left\{ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \left[(1-\vartheta)\bar{\psi} + \vartheta\right] (\beta^2 + \beta + 1)\beta\vartheta \right\} \\ &\quad - \left\{ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) + (1-\bar{\theta})\vartheta \right\} \vartheta \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta \\ &= \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta\vartheta \left\{ (1-\bar{\theta})\vartheta(\beta^2 + \beta + 1) - (1-\bar{\theta})\vartheta \right\}\end{aligned}$$

Finally, this yields

$$[\text{l.h.s.}] = \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta + 1)(1-\bar{\theta})\beta^2\vartheta^2. \tag{B.3.4}$$

Turning to the term in brackets on the right hand side of (B.3.2) one can split the sum in the first $\{\circ\}$-bracket and the first line at $(1-\vartheta)$ and re-order everything to get

$$\begin{aligned}[\text{r.h.s.}] = &\quad \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta \\ &- \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta\vartheta \\ &+ (1-\bar{\theta})\beta\vartheta(1-\vartheta) \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] + (1-\bar{\theta})(\beta + 1)\vartheta \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \beta\vartheta.\end{aligned}$$

The second and the fifth line can be aggregated into a new second line, while the third is split at the $(1-\bar{\theta})$-term

$$\begin{aligned}= &\quad \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1)(1-\bar{\theta})\beta\vartheta^2 \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})\beta\vartheta \\ &- \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})\beta\vartheta^2 \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\right] (1-\bar{\theta})(\beta + 1)\vartheta \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) \left[(1-\vartheta)\bar{\psi} + \vartheta\right].\end{aligned}$$

Aggregating the second and fourth line yields

$$\begin{aligned}= &\quad \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})(\beta + 1)\beta^2\vartheta^2 \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})\beta\vartheta \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\right] (1-\bar{\theta})(\beta + 1)\vartheta \\ &+ \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta^2 + \beta + 1) \left[(1-\vartheta)\bar{\psi} + \vartheta\right].\end{aligned}$$

Splitting now the first line in the second factor and factorizing the last two lines gives

$$\begin{aligned}
[\text{r.h.s.}] = \; & \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right]^2 (\beta^2+\beta)\beta + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right]^2 \beta \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})(\beta+1)\beta^2\vartheta^2 \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})\beta\vartheta \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \left\{ \begin{array}{l} +\vartheta \; (\beta+1) + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta(\beta+1) \\ -\vartheta\bar{\theta}(\beta+1) + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \end{array} \right\},
\end{aligned}$$

where the second line of the $\{\circ\}$-bracket can be summarized to

$$\begin{aligned}
= \; & \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right]^2 (\beta+1)\beta^2 + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right]^2 \beta \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})(\beta+1)\beta^2\vartheta^2 \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})\beta\vartheta \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \left[\vartheta(\beta+1) - \vartheta\bar{\theta}\beta + (1-\vartheta)\bar{\psi}\right] \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta+1)\beta.
\end{aligned}$$

Taking the second term of the first into the third line, from which $[(1-\vartheta)\bar{\psi}+\vartheta]$ can then be extracted, de-factorizing, and simplifying yields

$$\begin{aligned}
[\text{r.h.s.}] = \; & \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right]^2 (\beta+1)\beta^2 \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})(\beta+1)\beta^2\vartheta^2 \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \left\{ \begin{array}{l} \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta \\ + \left[\vartheta(\beta+1) - \vartheta\bar{\theta}\beta + (1-\vartheta)\bar{\psi}\right] \end{array} \right\} \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta+1)\beta\,. \qquad \text{(B.3.5)}
\end{aligned}$$

Because the new term in $\{\circ\}$ can be aggregated to $\left[(1-\vartheta)\bar{\psi} + \vartheta\right](\beta+1)$, one has finally for the right hand side

$$\begin{aligned}
[\text{r.h.s.}] = \; & \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right]^2 (\beta+1)\beta^2 \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (1-\bar{\theta})(\beta+1)\beta^2\vartheta^2 \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \left[(1-\vartheta)\bar{\psi} + \vartheta\right] (\beta+1) \\
& + \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta+1)\beta\,. \qquad \text{(B.3.6)}
\end{aligned}$$

All this allows to extract a *common factor* $\beta+1$ from the [r.h.s.] in (B.3.6). With a constant Ξ_1, defined as

$$\Xi_1 \equiv \left\{ \begin{array}{l} \left[(1-\vartheta)\bar{\psi} + \vartheta\right]^2 \\ + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \beta^2\vartheta^2(1-\bar{\theta}) \\ + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \left[(1-\vartheta)\bar{\psi} + \vartheta\right] \beta \\ + \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right]^2 \beta^2 \end{array} \right\}, \qquad \text{(B.3.7)}$$

one can finally rewrite the right hand side as

$$[\text{r.h.s.}] = (\beta + 1) \times \Xi_1 \tag{B.3.8}$$

Turning back to (B.3.3) and inserting (B.3.4) for the left and (B.3.8) for the right hand side yields

$$\left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] (\beta+1)(1-\bar{\theta})\beta^2\vartheta^2 = (\beta+1)\ \Xi_1\ \psi(t).$$

Solving this for $\psi(t)$ gives the desired (3.6.10) from the main text:

$$\psi(t) = \left[(1-\vartheta)\bar{\psi} + \vartheta\bar{\theta}\right] \frac{(1-\bar{\theta})\beta^2\vartheta^2}{\Xi_1}\,. \tag{B.3.9}$$

Alternative Values for β

Figure B.1 with different values for β complements Fig. 3.3 from the text, where the subjective time preference factor had been set at $\beta = 0.6$.
The effect of alternative values for the subjective time discount factor follows intuition: the higher β the more patient are the agents, and the less pronounced is their preference for present consumption. Consequently, agents save more which translates in higher values for Ψ, regardless of the parameters ϑ and $\bar{\theta}$.

Limits of Ψ

Due to (3.6.15) the limits of (3.6.14) are also the limits of Ψ. Therefore,

$$\lim_{\bar{\theta}\to 1} \Psi = \lim_{\bar{\theta}\to 1} \left\{ \frac{1}{\frac{[(1-\vartheta)\psi+\vartheta]^2}{[(1-\vartheta)\psi+\vartheta\bar{\theta}](1-\bar{\theta})\beta^2\vartheta^2} + \frac{[(1-\vartheta)\psi+\vartheta]}{(1-\bar{\theta})\beta\vartheta^2} + \frac{[(1-\vartheta)\psi+\vartheta\bar{\theta}]}{(1-\bar{\theta})\vartheta^2} + 1} \right\} = 0, \tag{B.3.10}$$

as the first three terms in the denominator approaches infinity because of $\lim_{\bar{\theta}\to 1}(1-\bar{\theta}) = 0$. Apart from eyeball-testing Fig. 3.3 there is no proof for $\lim_{\bar{\theta}\to 0}\Psi = 0$. However, inspection of (3.6.14) indicates that in the required iterations the denominator's first fraction seems to approach infinity since its denominator approaches zero faster than its nominator.
The Lucas [1978]-case of (3.6.17) is obtained by rearranging (3.6.14):

$$\lim_{\vartheta\to 0} \Psi = \lim_{\vartheta\to 0} \left\{ \frac{[(1-\vartheta)\psi + \vartheta\bar{\theta}](1-\bar{\theta})\beta^2\vartheta^2}{\begin{array}{c}[(1-\vartheta)\psi+\vartheta]^2 + [(1-\vartheta)\psi+\vartheta][(1-\vartheta)\psi+\vartheta\bar{\theta}]\beta \\ +[(1-\vartheta)\psi+\vartheta\bar{\theta}]^2\beta^2 + [(1-\vartheta)\psi+\vartheta\bar{\theta}](1-\bar{\theta})\beta^2\vartheta^2\end{array}} \right\} = 0, \tag{B.3.11}$$

since there are elements of order ϑ^0 in the denominator but not in the nominator.
Equation (3.6.18) is obtained by plugging $\vartheta = 1$ into (3.6.14)

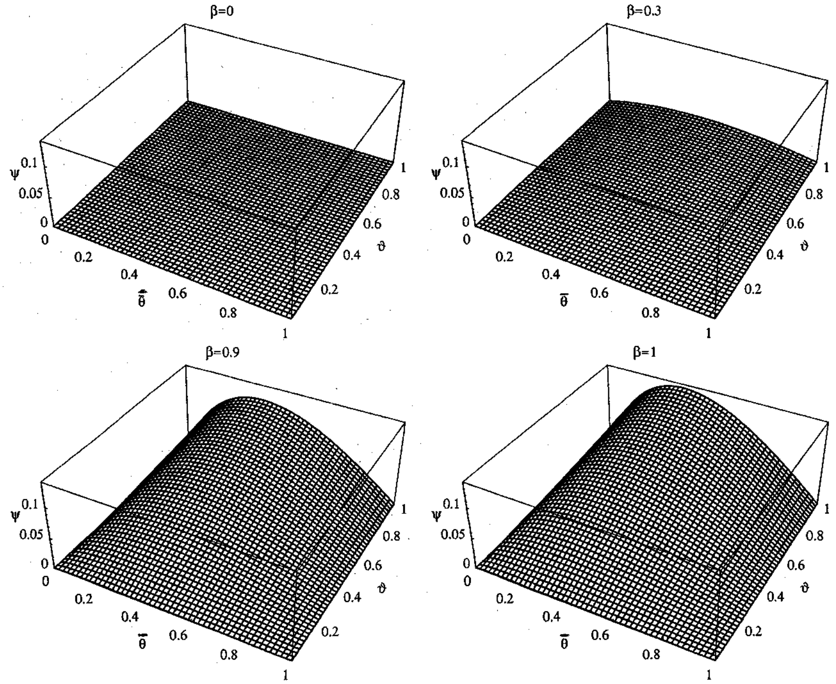

Fig. B.1. Investment-output ratio for alternative βs

Remarks: The plot visualizes the values of the investment-output ratio Ψ for different values of the technologies' elasticity parameters ϑ and $\bar{\theta}$ and a subjective discount factor of $\beta = 0$, $\beta = 0.3$, $\beta = 0.9$, and $\beta = 1$.

$$\begin{aligned} \Psi|_{\vartheta=1} &= \frac{[(1-1)\psi+\bar{\theta}](1-\bar{\theta})\beta^2 1^2}{[(1-1)\psi+1]^2+[(1-1)\psi+1][(1-1)\psi+\theta]\beta + [(1-1)\psi+\bar{\theta}]^2\beta^2+[(1-1)\psi+\bar{\theta}](1-\bar{\theta})\beta^2 1^2} \\ &= \frac{\beta^2\bar{\theta}(1-\bar{\theta})}{1+\beta\bar{\theta}+\beta^2\bar{\theta}} \,. \end{aligned} \tag{B.3.12}$$

The limit for infinite impatience, (3.6.19), follows from (3.6.14) by taking β^2 out of the denominator:

$$\lim_{\beta\to 0}\Psi = \lim_{\beta\to 0}\frac{\beta^2}{\frac{[(1-\vartheta)\psi+\vartheta]^2}{[(1-\vartheta)\psi+\vartheta\bar{\theta}](1-\bar{\theta})\vartheta^2}+\frac{[(1-\vartheta)\psi+\vartheta]}{(1-\bar{\theta})\vartheta^2}\beta+\frac{[(1-\vartheta)\psi+\vartheta\bar{\theta}]}{(1-\bar{\theta})\vartheta^2}\beta^2+\beta^2} = 0, \tag{B.3.13}$$

because the first term in the denominator does not contain any power of β.

B.3.3 Full Depreciation

In case of fully depreciating physical capital, equilibrium in the goods market is characterized by (3.6.32):

$$1 = \frac{1-\bar{\theta}}{(\beta^2+\beta+1)\bar{\theta}}\psi(t)|_{\vartheta=1} + \frac{\beta(\beta+1)\bar{\theta}+1}{(\beta^2+\beta+1)\beta\bar{\theta}}\psi(t)|_{\vartheta=1} + \frac{\beta(\beta+1)\bar{\theta}+1}{\beta^2+\beta+1} + \psi(t)|_{\vartheta=1} \,. \tag{B.3.14}$$

Solving for the investment-output ratio is straightforward:

$$1 - \frac{\beta(\beta+1)\bar{\theta}+1}{\beta^2+\beta+1} = \left[\frac{(1-\bar{\theta})\beta}{(\beta^2+\beta+1)\beta\bar{\theta}} + \frac{\beta(\beta+1)\bar{\theta}+1}{(\beta^2+\beta+1)\beta\bar{\theta}} + 1\right]\psi(t)|_{\vartheta=1}$$

After simplification one finally obtains

$$\psi(t)|_{\vartheta=1} = \frac{\beta^2\bar{\theta}(1-\bar{\theta})}{1+\beta\bar{\theta}+\beta^2\bar{\theta}} \,, \tag{B.3.15}$$

which proofs (3.6.33) in the text.

C

Replication with PAYGO

C.1 Replicating Tradability

C.1.1 Derivation of (4.4.14)

In order to derive the central condition (4.4.14) start with (4.4.13) from Sect. 4.4.2:

$$\left[\frac{(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'}{\vartheta'} + \frac{1-\bar{\theta}'}{\beta'^2+\beta'+1}\right]\frac{\tilde{\check{Y}}'(t+1)}{N_o'(t+1)} = \frac{1}{1+\tau_C}\times$$
$$\left\{\begin{array}{l}\left[\begin{array}{l}\dfrac{i_y'(t-1)}{N_m'(t)}\left(\theta'(t)\check{Y}'(t) + \dfrac{1-\vartheta'}{\vartheta'}\Psi'\check{Y}'(t)\right) + (1-\tau_\omega)(1-\theta'(t))\dfrac{\check{Y}'(t)}{N_m'(t)} \\ -(1+\tau_C)\left[\dfrac{1}{\vartheta'} + \dfrac{1-\bar{\theta}'}{(\beta'^2+\beta'+1)\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right]}\right]\dfrac{\Psi'}{\beta'}\dfrac{\check{Y}'(t)}{N_m'(t)}\end{array}\right] \\ \times\dfrac{\tilde{\check{Y}}'(t+1)}{\Psi'\check{Y}'(t)}\left(\vartheta'\tilde{\theta}'(t+1) + (1-\vartheta')\Psi'\right) + \dfrac{\tau_\omega(1-\tilde{\theta}'(t+1))\tilde{\check{Y}}'(t+1)}{N_o'(t+1)}\end{array}\right\}. \tag{C.1.1}$$

Due to (4.2.3) one can substitute $N_m'(t)$ for $N_o'(t+1)$ and cancel them as well as $\check{Y}'(t)$ and $\tilde{\check{Y}}'(t+1)$. Eliminating the fractions and bringing everything on one side yields

$$\begin{aligned}0 = {} & i_y'(t-1)\beta'\left[\vartheta'\theta'(t) + (1-\vartheta')\Psi'\right]\left[\vartheta'\tilde{\theta}'(t+1) + (1-\vartheta')\Psi'\right]\Xi_2 \\ & +(1-\tau_\omega)\beta'\vartheta'(1-\theta'(t))\left[\vartheta'\tilde{\theta}'(t+1) + (1-\vartheta')\Psi'\right]\Xi_2 \\ & -(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]\left[\vartheta'\tilde{\theta}'(t+1) + (1-\vartheta')\Psi'\right] \\ & +\tau_\omega\Psi'\beta'\vartheta'(1-\tilde{\theta}'(t+1))\Xi_2 \\ & -(1+\tau_C)\Psi'\beta'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right],\end{aligned}$$

where Ξ_2 is shorthand for the constant term of

$$\Xi_2 \equiv \left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right](\beta'^2 + \beta' + 1). \tag{C.1.2}$$

Due to $\beta \geq 0$, $\theta \geq 0$, $1 \geq \vartheta \geq 0$ and $\Psi \geq 0$ this constant is positive. Extension of all the relevant products, and reorganizing terms by factors of $\theta'(t)$ and $\tilde{\theta}'(t+1)$ allows then to establish

$$\begin{aligned}
0 = {} & \theta'(t)\,\tilde{\theta}'(t+1)\,\beta'\vartheta'^2\Xi_2\Big\{i'_y(t-1) - (1-\tau_\omega)\Big\} \\
& + \theta'(t)\,\beta'\vartheta'(1-\vartheta')\Psi'\Xi_2\Big\{i'_y(t-1) - (1-\tau_\omega)\Big\} \\
& + \tilde{\theta}'(t+1)\,\vartheta'\left\{\begin{array}{l} i'_y(t-1)\beta'(1-\vartheta')\Psi'\Xi_2 + (1-\tau_\omega)\beta'\vartheta'\Xi_2 \\ -(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right] - \tau_\omega\Psi'\beta'\Xi_2 \end{array}\right\} \\
& + \Psi'\left\{\begin{array}{l} +\, i'_y(t-1)\beta'(1-\vartheta')^2\Psi'\Xi_2 \\ -(1+\tau_C)\beta'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] \\ -(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right](1-\vartheta') \\ +\, \beta'\vartheta'(1-\vartheta')\Xi_2 + \tau_\omega\beta'\vartheta'^2\Xi_2 \end{array}\right\}.
\end{aligned} \tag{C.1.3}$$

This is (4.4.14) in the text.

C.1.2 Derivation of Public Sector Parameters

Derivation of the public sector parameters is based on the three conditions (4.4.15) to (4.4.17) which are repeated here for convenience:

$$0 = i'_y(t-1) - (1-\tau_\omega), \tag{C.1.4}$$

$$0 = \left\{\begin{array}{l} i'_y(t-1)\beta'(1-\vartheta')\Psi'\Xi_2 + (1-\tau_\omega)\beta'\vartheta'\Xi_2 \\ -(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right] - \tau_\omega\Psi'\beta'\Xi_2 \end{array}\right\}, \tag{C.1.5}$$

$$\text{and}\quad 0 = \left\{\begin{array}{l} +\, i'_y(t-1)\beta'(1-\vartheta')^2\Psi'\Xi_2 \\ -(1+\tau_C)\beta'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] \\ -(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right](1-\vartheta') \\ +\, \beta'\vartheta'(1-\vartheta')\Xi_2 + \tau_\omega\beta'\vartheta'^2\Xi_2 \end{array}\right\}. \tag{C.1.6}$$

Derivation of Contribution Rate

From (C.1.4) it follows that

$$i'_y(t-1) = 1 - \tau_\omega. \tag{C.1.7}$$

As τ_ω is constant over time so must the investment fraction; i.e. it must hold that

$$i'_y = i'_y(t-1) = i'_y(t) = 1 - \tau_\omega. \tag{C.1.8}$$

Using this one can write (C.1.5) as

$$\begin{aligned}&(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right] \\ &= \beta'\Xi_2\left[(1-\vartheta')\Psi' + \vartheta'\right] - \tau_\omega\beta'\Xi_2\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right].\end{aligned} \tag{C.1.9}$$

Substituting (C.1.8) for $i'_y(t-1)$ in (C.1.6) yields after multiplication of both sides with Ψ':

$$\begin{aligned}&(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]\left\{\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] + \Psi'(1-\vartheta')\right\} \\ &= (1-\tau_\omega)\beta'(1-\vartheta')^2\Psi'^2\Xi_2 + \beta'\vartheta'(1-\vartheta')\Psi'\Xi_2 + \tau_\omega\beta'\vartheta'^2\Psi'\Xi_2.\end{aligned} \tag{C.1.10}$$

Note that τ_C can be eliminated from (C.1.10) by replacing the first three factors on its left hand side with the right hand side of (C.1.9). After canceling β's and Ξ_2s this reads

$$\begin{aligned}&\left\{\begin{array}{l}\left[(1-\vartheta')\Psi' + \vartheta'\right] \\ -\tau_\omega\left(2\Psi' - \vartheta'\Psi' + \vartheta'\right)\end{array}\right\}\left\{\begin{array}{l}\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] \\ +\Psi'(1-\vartheta')\end{array}\right\} \\ &= (1-\tau_\omega)(1-\vartheta')^2\Psi'^2 + \vartheta'(1-\vartheta')\Psi' + \tau_\omega\vartheta'^2\Psi'.\end{aligned} \tag{C.1.11}$$

Solving (C.1.11) for τ_ω gives

$$\tau_\omega = \frac{\left[(1-\vartheta')\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right]}{\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] + \Psi'(\Psi' + \vartheta' - 3\vartheta'\Psi' + 2\vartheta'^2\Psi')}. \tag{C.1.12}$$

This is (4.4.18) in the text.

Derivation of Tax Rate

Use (C.1.9) and plug in (C.1.12) for τ_ω to get an equation only in τ_C:

$$\begin{aligned}&(1+\tau_C)\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right] \\ &= \beta'\Xi_2\left[(1-\vartheta')\Psi' + \vartheta'\right] - \beta'\Xi_2\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right] \\ &\quad\times\frac{\left[(1-\vartheta')\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right]}{\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] + \Psi'(\Psi' + \vartheta' - 3\vartheta'\Psi' + 2\vartheta'^2\Psi')}.\end{aligned}$$

Thus the replicating tax rate is given by

$$\begin{aligned}\tau_C = {}& \frac{\beta'\Xi_2\left[(1-\vartheta')\Psi' + \vartheta'\right]}{\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]} - 1 - \frac{1}{\Psi'\left[\Xi_2 + \vartheta'(1-\bar{\theta}')\right]}\times \\ & \frac{\beta'\Xi_2\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right]\left[(1-\vartheta')\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right]}{\left[2\Psi' - \vartheta'\Psi' + \vartheta'\right]\beta'\left[(1-\vartheta')\Psi' + \vartheta'\bar{\theta}'\right] + \Psi'(\Psi' + \vartheta' - 3\vartheta'\Psi' + 2\vartheta'^2\Psi')}.\end{aligned} \tag{C.1.13}$$

As the first two terms can be summarized to

$$\frac{\beta'\Xi_2\left[(1-\vartheta')\Psi'+\vartheta'\right]}{\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]}-1=\frac{\beta'\Xi_2\left[(1-\vartheta')\Psi'+\vartheta'\right]-\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]}{\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]}$$
$$=\frac{\left[\beta'\Psi'\Xi_2-\beta'\vartheta'\Psi'\Xi_2+\beta'\vartheta'\Xi_2\right]-\left[\Psi'\Xi_2+\vartheta'\Psi'-\bar{\theta}'\vartheta'\Psi'\right]}{\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]}$$

the replicating tax rate is given by

$$\tau_C=\frac{\left[\beta'\Psi'\Xi_2-\beta'\vartheta'\Psi'\Xi_2+\beta'\vartheta'\Xi_2\right]-\left[\Psi'\Xi_2+\vartheta'\Psi'-\bar{\theta}'\vartheta'\Psi'\right]}{\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]}$$
$$-\frac{1}{\Psi'\left[\Xi_2+\vartheta'(1-\bar{\theta}')\right]}\times$$
$$\frac{\beta'\Xi_2\left[2\Psi'-\vartheta'\Psi'+\vartheta'\right]\left[(1-\vartheta')\Psi'+\vartheta'\right]\beta'\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]}{\left[2\Psi'-\vartheta'\Psi'+\vartheta'\right]\beta'\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]+\Psi'(\Psi'+\vartheta'-3\vartheta'\Psi'+2\vartheta'^2\Psi')}.\tag{C.1.14}$$

which is (4.4.19).

Derivation of Investment Fraction

Finally, using the solution for τ_ω, (C.1.12), in (C.1.8) gives

$$i_y'=\frac{\Psi'\beta'\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]+\Psi'(\Psi'+\vartheta'-3\vartheta'\Psi'+2\vartheta'^2\Psi')}{\left[2\Psi'-\vartheta'\Psi'+\vartheta'\right]\beta'\left[(1-\vartheta')\Psi'+\vartheta'\bar{\theta}'\right]+\Psi'(\Psi'+\vartheta'-3\vartheta'\Psi'+2\vartheta'^2\Psi')},\tag{C.1.15}$$

which proofs (4.4.20).

C.2 Full Depreciation

C.2.1 Replicating Parameters

Based on the results of Chapt. 3, (3.6.33) and (3.6.34), one has

$$\Psi'|_{\vartheta'=1}=\psi'(t)|_{\vartheta=1}=\frac{\beta'^2\bar{\theta}'(1-\bar{\theta}')}{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}\tag{C.2.1}$$

for the investment-output ratio in the case of fully depreciating physical capital. Furthermore, for $\vartheta'=1$ the auxiliary constant Ξ_2 can be reduced from (C.1.2) to

$$\Xi_2|_{\vartheta'=1}=\left[(1-1)\Psi'+1\times\bar{\theta}'\right](\beta'^2+\beta'+1)=\bar{\theta}'(\beta'^2+\beta'+1).\tag{C.2.2}$$

Contribution Rate

By using $\vartheta' = 1$ in (C.1.12) this simplifies to

$$\tau_\omega|_{\vartheta'=1} = \frac{[(1-1)\Psi' + 1]\,\beta'\left[(1-1)\Psi' + 1\times\bar{\theta}'\right]}{[2\Psi' - \Psi' + 1]\,\beta'\left[(1-1)\Psi' + 1\times\bar{\theta}'\right] + \Psi'(\Psi' + 1 - 3\Psi' + 2\Psi')}$$
$$= \frac{\beta'\bar{\theta}'}{[\Psi' + 1]\,\beta'\bar{\theta}' + \Psi'}. \tag{C.2.3}$$

Applying now (C.2.1) in this reads

$$\tau_\omega|_{\vartheta'=1} = \frac{\beta'\bar{\theta}'}{\left[\frac{\beta'^2\bar{\theta}'(1-\bar{\theta}')}{1+\beta'\bar{\theta}+\beta'^2\bar{\theta}'} + 1\right]\beta'\bar{\theta}' + \frac{\beta'^2\bar{\theta}'(1-\bar{\theta}')}{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}}$$
$$= \frac{1 + \beta'\bar{\theta}' + \beta'^2\bar{\theta}'}{1 + \beta' + 2\beta'^2\bar{\theta}' - \beta'^2\bar{\theta}'^2}, \tag{C.2.4}$$

which is (4.4.21) in the main text.

Tax Rate

Inserting $\vartheta' = 1$ into (C.1.13) allows to simplify the solution to

$$\tau_C|_{\vartheta'=1}$$
$$= \frac{\beta'\Xi_2\,[(1-1)\Psi' + 1]}{\Psi'\left[\Xi_2 + 1\times(1-\bar{\theta}')\right]} - 1 - \frac{1}{\Psi'\left[\Xi_2 + 1(1-\bar{\theta}')\right]}\times$$
$$\frac{\beta'\Xi_2\,[2\Psi' - 1\times\Psi' + 1]\,[(1-1)\Psi' + 1]\,\beta'\left[(1-1)\Psi' + 1\times\bar{\theta}'\right]}{[2\Psi' - 1\times\Psi' + 1\times]\,\beta'\left[(1-1)\Psi' + 1\times\bar{\theta}'\right] + \Psi'(\Psi' + 1 - 3\Psi' + 2\Psi')}$$
$$= \frac{\beta'\Xi_2}{\left(\Xi_2 + 1 - \bar{\theta}'\right)\left[(\Psi' + 1)\,\beta'\bar{\theta}' + \Psi'\right]} - 1\,. \tag{C.2.5}$$

Applying then the results of (C.2.1) and (C.2.2) gives after some simplifications

$$\tau_C|_{\vartheta'=1} = \frac{\beta'^2(1-\bar{\theta}')^2}{1 + \beta' + 2\beta'^2\bar{\theta}' - \beta'^2\bar{\theta}'^2}. \tag{C.2.6}$$

This proofs (4.4.22) from the main text.

Investment Fraction

To derive the investment fraction in case of full depreciation, start with $\vartheta' = 1$ in (C.1.15) and simplify to obtain

$$i'_y|_{\vartheta'=1} = \frac{\Psi'\beta'\bar{\theta}' + \Psi'}{[\Psi' + 1]\,\beta'\bar{\theta}' + \Psi'}.$$

Applying again (C.2.1) for Ψ' allows to reduce this finally to

$$i'_y|_{\vartheta'=1} = \frac{\beta'(1-\bar{\theta}'+\beta'\bar{\theta}'-\beta'\bar{\theta}'^2)}{1+\beta'+2\beta'^2\bar{\theta}'-\beta'^2\bar{\theta}'^2}. \tag{C.2.7}$$

This is the main text's (4.4.23).

C.2.2 Alternative Derivation

Derivation of Equation System

To derive the equation system (4.4.30) to (4.4.32) take (4.4.29),

$$\begin{aligned}\frac{\beta'(\beta'+1)\bar{\theta}'+1}{\beta'^2+\beta'+1}\frac{Y'(t+1)}{N'_o(t+1)} &= \frac{1}{1+\tau_C|_{\vartheta'=1}}\\ &\times\left\{\begin{array}{l}\left[\begin{array}{l}\frac{i'_y(t-1)|_{\vartheta'=1}}{N'_m(t)}\theta'(t)Y'(t)+(1-\tau_\omega|_{\vartheta'=1})(1-\theta'(t))\frac{Y'(t)}{N'_m(t)}\\ -(1+\tau_C|_{\vartheta'=1})\frac{\beta'(\beta'+1)\bar{\theta}'+1}{(\beta'^2+\beta'+1)\beta'\bar{\theta}'}\frac{I'(t)}{N'_m(t)}\end{array}\right]\\ \times\frac{\tilde{\theta}'(t+1)\tilde{Y}'(t+1)}{I'(t)}+\frac{\tau_\omega|_{\vartheta'=1}(1-\tilde{\theta}'(t+1))\tilde{Y}'(t+1)}{N'_o(t+1)}\end{array}\right\},\end{aligned} \tag{C.2.8}$$

cancel $N'_m(t) = N'_o(t+1)$ as well as $\tilde{Y}'(t+1)$, and replace $I'(t)$ by $Y'(t) \times \Psi'|_{\vartheta'=1}$ to get

$$\begin{aligned}&(1+\tau_C|_{\vartheta'=1})\frac{\beta'(\beta'+1)\bar{\theta}'+1}{\beta'^2+\beta'+1}\\ &=\left\{\begin{array}{l}i'_y(t-1)|_{\vartheta'=1}\theta'(t)\tilde{\theta}'(t+1)\frac{1}{\Psi'|_{\vartheta'=1}}\\ +(1-\tau_\omega|_{\vartheta'=1})(1-\theta'(t))\tilde{\theta}'(t+1)\frac{1}{\Psi'|_{\vartheta'=1}}\\ -(1+\tau_C|_{\vartheta'=1})\frac{\beta'(\beta'+1)\bar{\theta}'+1}{(\beta'^2+\beta'+1)\beta'\bar{\theta}'}\tilde{\theta}'(t+1)\\ +\tau_\omega|_{\vartheta'=1}(1-\tilde{\theta}'(t+1))\end{array}\right\}.\end{aligned} \tag{C.2.9}$$

Plugging in $\Psi'|_{\vartheta'=1} = \frac{\beta'^2\bar{\theta}'(1-\bar{\theta}')}{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}$ from (C.2.1) allows to restate this as

$$0 = \tilde{\theta}'(t)\tilde{\theta}'(t+1)(\beta'^2+\beta'+1)(1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}')\begin{bmatrix} i'_y(t-1)|_{\vartheta'=1} \\ -(1-\tau_\omega|_{\vartheta'=1}) \end{bmatrix}$$
$$+\tilde{\theta}'(t+1)(\beta'^2+\beta'+1)\begin{bmatrix} (1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}') \\ -\tau_\omega|_{\vartheta'=1}(1+\beta'\bar{\theta}'+2\beta'^2\bar{\theta}'-\beta'^2\bar{\theta}'^2) \\ -\beta'(1-\bar{\theta}')\dfrac{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}{\beta'^2+\beta'+1} \\ -\tau_C|_{\vartheta'=1}\beta'(1-\bar{\theta}')\dfrac{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}{\beta'^2+\beta'+1} \end{bmatrix}$$
$$+\beta'^2(1-\bar{\theta}')\bar{\theta}'\left[\tau_\omega|_{\vartheta'=1}(\beta'^2+\beta'+1)-(1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}')(1+\tau_C|_{\vartheta'=1})\right]. \tag{C.2.10}$$

This condition will be satisfied for *any* realizations of the stochastics $\tilde{\theta}'(t)$ and $\tilde{\theta}'(t+1)$ if

$$0 = i'_y(t-1)|_{\vartheta'=1} - (1-\tau_\omega|_{\vartheta'=1}), \tag{C.2.11}$$
$$0 = (1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}') - \tau_\omega|_{\vartheta'=1}(1+\beta'\bar{\theta}'+2\beta'^2\bar{\theta}'-\beta'^2\bar{\theta}'^2)$$
$$-\beta'(1-\bar{\theta}')\frac{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}{\beta'^2+\beta'+1} - \tau_C|_{\vartheta'=1}\beta'(1-\bar{\theta}')\frac{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}{\beta'^2+\beta'+1}, \tag{C.2.12}$$

and $$0 = \tau_\omega|_{\vartheta'=1}(\beta'^2+\beta'+1) - (1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}')(1+\tau_C|_{\vartheta'=1}). \tag{C.2.13}$$

These are the conditions (4.4.30) to (4.4.32).

Solving the System

To solve above system of equations for $\tau_\omega|_{\vartheta'=1}$, $\tau_C|_{\vartheta'=1}$ and $i'_y(t-1)|_{\vartheta'=1}$ take the implication of the third,

$$\tau_\omega|_{\vartheta'=1} = (1+\tau_C|_{\vartheta'=1})\frac{1+\beta\bar{\theta}+\beta^2\bar{\theta}}{\beta^2+\beta+1}, \tag{C.2.14}$$

and use this in (C.2.12) to get an equation only in $\tau_C|_{\vartheta'=1}$. Solving above equation gives the solution

$$\tau_C|_{\vartheta'=1} = \frac{\beta'^2(1-\bar{\theta}')^2}{1+\beta'+2\beta'^2\bar{\theta}'-\beta'^2\bar{\theta}'^2}. \tag{C.2.15}$$

Using this result in (C.2.14) allows to specify the solution of the contribution rate as

$$\tau_\omega|_{\vartheta'=1} = \frac{1+\beta'\bar{\theta}'+\beta'^2\bar{\theta}'}{1+\beta'+2\beta'^2\bar{\theta}'-\beta'^2\bar{\theta}'^2}. \tag{C.2.16}$$

Finally, from (C.2.11) one gets

$$i'_y(t-1)|_{\vartheta'=1} = \frac{\beta'(1-\bar{\theta}' + \beta'\bar{\theta}' - \beta'\bar{\theta}'^2)}{1+\beta' + 2\beta'^2\bar{\theta}' - \beta'^2\bar{\theta}'^2}. \tag{C.2.17}$$

As this solution is independent of the time period t, one can write $i'_y(t-1)|_{\vartheta'=1} = i'_y|_{\vartheta'=1}$. Since (C.2.15), (C.2.16) and (C.2.17) are then identical to (4.4.21), (4.4.22) and (4.4.23) this concludes the alternative derivation.

D

Discussion and Assessment

D.1 Analyzing the Framework

Human Capital Securities

To show that the human capital security fits well into asset pricing's standard concept of stochastic discount factors, one inserts (3.3.9) and (3.5.29) into the left hand side of the Euler equation (3.5.28):

$$\begin{aligned}
&\mathbb{E}_t\left[\tilde{M}_i(t+1)\times\tilde{R}_H(t+1)\right]\\
&=\mathbb{E}_t\left[\begin{array}{l}\dfrac{1}{(1-\vartheta)\Psi+\vartheta\bar{\theta}}\,\dfrac{(1-\vartheta)\Psi+\vartheta}{(1-\vartheta)\Psi+\vartheta}\,\dfrac{I(t)}{\tilde{Y}(t+1)}\\[2ex] \times\left[(1-\vartheta)\Psi+\vartheta\bar{\theta}\right]\dfrac{1-\tilde{\theta}(t+1)}{1-\bar{\theta}}\,\dfrac{\tilde{Y}(t+1)}{I(t)}\end{array}\right]
\end{aligned}$$

and simplifies

$$=\mathbb{E}_t\left[\frac{1-\tilde{\theta}(t+1)}{1-\bar{\theta}}\right]=\frac{1-\mathbb{E}_t[\tilde{\theta}(t+1)]}{1-\bar{\theta}}=1. \tag{D.1.1}$$

Since this yield one, the central condition is fulfilled for the human capital security.

Riskless Security

A present value statement cannot be cross checked, as no assumptions on the price of a bond with a payoff corresponding to the right hand side of (5.2.5) have been made. Naturally, it would be given by $1/R(t)$, because all returns are gross returns. However, one can rewrite (5.2.5) as

$$Y(t)=\frac{(1-\vartheta)\Psi+\vartheta\bar{\theta}}{\Psi R(t)}\left(\mathbb{E}_t\left[\frac{1}{\tilde{Y}(t+1)}\right]\right)^{-1}, \tag{D.1.2}$$

which is essentially a present value statement of future output. It crucially rests on the simple connection of investment and output as implied by the simple savings policies. Still, (D.1.2) could be used to investigate long-run macroeconomic outlook, because $Y(t)$ as well as $R(t)$ are observable at t.

D.2 Technological Progress

D.2.1 Stochastic Growth Theory

Applying Itô's lemma to the definition of k_t gives:

$$\begin{aligned} dk_t = & \frac{\partial k_t}{\partial A_t} dA_t + \frac{\partial k_t}{\partial L_t} dL_t + \frac{\partial k_t}{\partial K_t} dK_t \\ & + \frac{1}{2} \left[\frac{\partial^2 k_t}{\partial A_t^2} (dA_t)^2 + \frac{\partial^2 k_t}{\partial L_t^2} (dL_t)^2 + \frac{\partial^2 k_t}{\partial K_t^2} (dK_t)^2 \right] \\ & + \frac{\partial^2 k_t}{\partial A_t \partial L_t} dA_t dL_t + \frac{\partial^2 k_t}{\partial A_t \partial K_t} dA_t dK_t + \frac{\partial^2 k_t}{\partial L_t \partial K_t} dL_t dK_t. \end{aligned} \tag{D.2.1}$$

The necessary partial derivatives are

$$\begin{aligned} & \frac{\partial k_t}{\partial A_t} = \frac{-K_t}{A_t^2 L_t} = -\frac{K_t}{A_t} \quad \frac{\partial^2 k_t}{\partial A_t^2} = \frac{2K_t}{A_t^3 L_t} = \frac{2K_t}{A_t^2} \quad \frac{\partial^2 k_t}{\partial A_t \partial L_t} = \frac{K_t}{A_t^2 L_t^2} = \frac{K_t}{A_t L_t} \\ & \frac{\partial k_t}{\partial L_t} = \frac{-K_t}{A_t L_t^2} = -\frac{K_t}{L} \quad \frac{\partial^2 k_t}{\partial L_t^2} = \frac{2K_t}{A_t L_t^3} = \frac{2K_t}{L_t^2} \quad \frac{\partial^2 k_t}{\partial L_t \partial K_t} = \frac{-1}{A_t L_t^2} \\ & \frac{\partial k_t}{\partial K_t} = \frac{1}{A_t L_t} \quad \frac{\partial^2 k_t}{\partial K_t^2} = 0 \quad \frac{\partial^2 k_t}{\partial A_t \partial K_t} = \frac{-1}{A_t^2 L_t} \end{aligned} \tag{D.2.2}$$

Furthermore, the rules of stochastic calculus imply for the processes (5.3.4) to (5.3.6):

$$\begin{aligned} & (dA_t)^2 = A_t^2 \sigma_A^2 dt + o(dt) \qquad dA_t dL_t = A_t L_t \sigma_A \sigma_L \rho_{z_A, z_L} dt + o(dt) \\ & (dL_t)^2 = L_t^2 \sigma_L^2 dt + o(dt) \qquad dA_t dK_t = o(dt) \\ & (dk_t)^2 = o(dt) \qquad dL_t dK_t = o(dt) \end{aligned} \tag{D.2.3}$$

Putting all this into (D.2.1) and simplifying with $\sigma_{AL} \equiv \sigma_A \sigma_L \rho_{z_L, z_A}$ and $\sigma_{A+L}^2 \equiv \sigma_A^2 + \sigma_L^2 + 2\sigma_{AL}$ yields

$$\begin{aligned} dk_t = & -\frac{k_t}{A_t} (\alpha A_t dt + \sigma_A A_t dz_A) - \frac{k_t}{L_t} (n L_t dt + \sigma_L L_t dz_L) + \frac{(sY_t dt - \delta K_t dt)}{A_t L_t} \\ & + \frac{1}{2} \left[\frac{2K_t}{A_t^2} A_t^2 \sigma_A^2 dt + \frac{2K_t}{L_t^2} L_t^2 \sigma_L^2 dt \right] + \frac{k_t}{A_t L_t} A_t L_t \sigma_A \sigma_L \rho_{z_A, z_L} dt \\ = & \left[sf(k_t) - k_t(\alpha + n + \delta) + k_t(\sigma_{A+L}^2 - \sigma_{AL}) \right] dt - k_t (\sigma_A dz_1 + \sigma_L dz_2), \end{aligned} \tag{D.2.4}$$

which is (5.3.7) in the main text.

D.2.2 Derivation of Stationary Distribution

In order to show that the process of k_t converges asymptotically to a stationary distribution it is sufficient to extend the proof of Merton [1975] to the case *with* technological progress as specified in (5.3.5). Following his approach, one can rewrite equation (5.3.7) by conveniently extending the analytical derivations and definitions as

$$dk_t = \left[sf(k_t) - k_t(\alpha + n + \delta - \sigma_\epsilon^2)\right] dt - k_t\sigma_k dZ, \tag{D.2.5}$$

where $\sigma_\epsilon^2 = \sigma_{A+L}^2 - \sigma_{AL} = \sigma_A^2 + \sigma_L^2 + \sigma_A\sigma_L\rho_{z_A,z_L}$ and the two normally distributed dz_A and dz_L have been regrouped into one equally normal distributed increment by applying the conventional variance-covariance rules:[1]

$$\sigma_A dz_A + \sigma_L dz_L = \sigma_{A+L} dZ. \tag{D.2.6}$$

Equation (D.2.5) is structurally identical to equation (9) in Merton [1975], so his analysis of the stationary distribution of k_t can directly be extended to the case presented here. Since neither existence nor uniqueness properties are the major goal here, one can directly assume that the required sufficient conditions are given: $s(k_t) > 0 \; \forall \; k_t < \overline{k}$ for some $\overline{k} > 0$ and $\alpha + n + \delta - \sigma_\epsilon^2 > 0$.[2] Let $\pi(k_t)$ be then the density function of the steady-state probability distribution for k_t. As shown in Malliaris and Brock [1982, Chap. 3.5], Bourguignon [1974, Sect. 3] or Merton [1975, App. B] the density function $\pi(x)$ of a diffusion process with stochastic differential representation of the form $dx = b(x)dt - \sqrt{a(x)}dZ$ is given by:

$$\pi(x) = \frac{\mathfrak{m}}{\mathfrak{a}(x)} \exp\left[\int^x \frac{2\mathfrak{b}(u)}{\mathfrak{a}(u)} du\right] \tag{D.2.7}$$

where $\mathfrak{m}$ is chosen so that $\int_0^\infty \pi(u)du = 1$. The coefficients for the density function in the case of (5.3.7) are

$$\mathfrak{a}(k_t) = \sigma_{A+L}^2 k_t^2 \tag{D.2.8}$$

$$\text{and} \quad \mathfrak{b}(k_t) = sf(k_t) - k_t(\alpha + n + \delta - \sigma_\epsilon^2) \tag{D.2.9}$$

Using these specifications in (D.2.7) gives

$$\begin{aligned}
\pi(k) &= \frac{\mathfrak{m}}{\sigma_{A+L}^2 k_t^2} \exp\left[2\int^{k_t} \frac{sf(u) - u(\alpha + n + \delta - \sigma_\epsilon^2)}{\sigma_{A+L}^2 u^2} du\right] \\
&= \frac{\mathfrak{m}}{\sigma_{A+L}^2 k_t^2} \exp\left[-2\frac{(\alpha + n + \delta - \sigma_\epsilon^2)}{\sigma_{A+L}^2} \ln k_t\right] \exp\left[\frac{2}{\sigma_{A+L}^2}\int^{k_t} \frac{sf(u)}{u^2} du\right] \\
&= \frac{\mathfrak{m}}{\sigma_{A+L}^2} k_t^{-2\frac{\alpha+n+\delta+\sigma_{AL}}{\sigma_{A+L}^2}} \exp\left[\frac{2}{\sigma_{A+L}^2}\int^{k_t} \frac{sf(u)}{u^2} du\right],
\end{aligned} \tag{D.2.10}$$

[1] This draws on Bourguignon [1974].

[2] Otherwise the drift in equation (5.3.7) would grow exponentially in time. Furthermore, the third condition – that $f(k_t)$ is concave and satisfies the well-known Inada conditions is already implied by the setting in Sect. 5.3.1.

where the last line follows from the definition of σ_ϵ^2. Since this is a logical extension of the earlier work on stochastic growth theory, where only m must be adjusted, asymptotic convergence of the distribution is given as in Merton [1975].

D.2.3 Cobb-Douglas Case

Rentals and Wages

Application of Itô's lemma on the formulation for r_t and ω_t as given in equations (5.3.22) and (5.3.23) reads:

$$dr_t = \frac{\partial r_t}{\partial k_t} dk_t + \frac{1}{2}\frac{\partial^2 r_t}{\partial k_t^2}(dk_t)^2 \tag{D.2.11}$$

$$\begin{aligned} d\omega_t = {} & \frac{\partial \omega_t}{\partial A_t} dA_t + \frac{\partial \omega_t}{\partial k_t} dk_t + \frac{1}{2}\left[\frac{\partial^2 \omega_t}{\partial A_t^2}(dA_t)^2 + \frac{\partial^2 \omega_t}{\partial \omega_t^2}(dk_t)^2\right] \\ & + \frac{\partial^2 \omega_t}{\partial A_t \partial k_t} dA_t dk_t \end{aligned} \tag{D.2.12}$$

Using the partial derivatives from $r_t = \theta k_t^{\theta-1}$ and $\omega_t = (1-\theta)A_t k_t^\theta$, namely

$$\begin{aligned} &\frac{\partial r_t}{\partial k_t} = \theta(\theta-1)k_t^{\theta-2} && \frac{\partial^2 r_t}{\partial k_t^2} = \theta(\theta-1)(\theta-2)k_t^{\theta-3} && \\ &\frac{\partial \omega_t}{\partial A_t} = (1-\theta)k_t^\theta && \frac{\partial^2 \omega_t}{\partial A_t \partial k_t} = \theta(1-\theta)k_t^{\theta-1} && \frac{\partial^2 \omega_t}{\partial A_t^2} = 0 \\ &\frac{\partial \omega_t}{\partial k_t} = \theta(1-\theta)A_t k_t^{\theta-1} && \frac{\partial^2 \omega_t}{\partial k_t^2} = -\theta(1-\theta)^2 A_t k_t^{\theta-2}, && \end{aligned} \tag{D.2.13}$$

as well as the results of stochastic calculus

$$\begin{aligned} (dA_t)^2 &= A_t^2\sigma_A^2 dt + o(dt) \\ (dk_t)^2 &= k_t^2\left[\sigma_A^2 + \sigma_L^2 + 2\sigma_A\ \sigma_L \rho_{z_A,z_L}\right] dt + o(dt) = k_t^2 \sigma_{A+L}^2 dt \\ dA_t dk_t &= A_t \sigma_A dz_A(-k)\left[\sigma_A dz_A + \sigma_L dz_L\right] + o(dt) = -A_t k_t\left[\sigma_A^2 + \sigma_{AL}\right] dt \end{aligned}$$

in (D.2.11) gives for the dynamics of rentals paid to physical capital:

$$dr_t = (\theta-1)r_t\left\{\left[sk_t^{\theta-1} - \alpha - n + \frac{\theta}{2}\sigma_{A+L}^2 - \sigma_{AL}\right] dt - (\sigma_A dz_A + \sigma_L dz_L)\right\}. \tag{D.2.14}$$

This is (5.3.24). The corresponding formulation for the wages is derived from

$$\begin{aligned}
d\omega_t &= (1-\theta)k_t^{\theta}\left(\alpha A_t dt + \sigma_A A_t dz_A\right) + \theta(1-\theta)A_t k_t^{\theta-1}k_t \\
&\quad \times\left\{\left[sk_t^{\theta-1} - (\alpha + n - \sigma_{A+L}^2 + \sigma_{AL})\right]dt - \left[\sigma_A dz_A + \sigma_L dz_L\right]\right\} \\
&\quad -\frac{1}{2}\theta(1-\theta)^2 A_t k_t^{\theta-2}k_t^2\sigma_{A+L}^2 dt + \theta(1-\theta)k_t^{\theta-1}(-A_t k_t)\left[\sigma_A^2 + \sigma_{AL}\right]dt \\
&= \omega_t\left\{\alpha dt + \theta\left[sk_t^{\theta-1} - (\alpha+n) - \frac{1}{2}(1-\theta)\sigma_{A+L}^2 + \sigma_L^2\right]dt\right. \\
&\quad \left. + \left[(1-\theta)\sigma_A dz_A - \theta\sigma_L dz_L\right]\right\}. \qquad \text{(D.2.15)}
\end{aligned}$$

This proofs the main text's (5.3.25).

Income from Human Labor and Physical Capital

Applying Itô's lemma on the definition of income paid to physical capital, $Y_{r,t} = r_t k_t A_t L_t$, yields

$$\begin{aligned}
dY_{r,t} &= \frac{\partial Y_{r,t}}{\partial r_t}dr_t + \frac{\partial Y_{r,t}}{\partial k_t}dk_t + \frac{\partial Y_{r,t}}{\partial A_t}dA_t + \frac{\partial Y_{r,t}}{\partial L_t}dL_t \\
&\quad + \frac{1}{2}\left[\frac{\partial^2 Y_{r,t}}{\partial r_t^2}(dr_t)^2 + \frac{\partial^2 Y_{r,t}}{\partial k_t^2}(dk_t)^2 + \frac{\partial^2 Y_{r,t}}{\partial A_t^2}(dA_t)^2 + \frac{\partial^2 Y_{r,t}}{\partial L_t^2}(dL_t)^2\right] \\
&\quad + \frac{\partial^2 Y_{r,t}}{\partial r_t \partial k_t}dr_t dk_t + \frac{\partial^2 Y_{r,t}}{\partial r_t \partial A_t}dr_t dA_t + \frac{\partial^2 Y_{r,t}}{\partial r_t \partial L_t}dr_t dL_t \\
&\quad + \frac{\partial^2 Y_{r,t}}{\partial k_t \partial A_t}dk_t dA_t + \frac{\partial^2 Y_{r,t}}{\partial k_t \partial L_t}dk_t dL_t + \frac{\partial^2 Y_{r,t}}{\partial A_t \partial L_t}dA_t dL_t. \qquad \text{(D.2.16)}
\end{aligned}$$

The required partial derivative are

$$\begin{aligned}
&\frac{\partial Y_{r,t}}{\partial r_t} = k_t A_t L_t && \frac{\partial^2 Y_{r,t}}{\partial r_t^2} = 0 && \frac{\partial^2 Y_{r,t}}{\partial r_t \partial k_t} = A_t L_t && \frac{\partial^2 Y_{r,t}}{\partial r_t \partial A_t} = k_t L_t \\
&\frac{\partial Y_{r,t}}{\partial k_t} = r_t A_t L_t && \frac{\partial^2 Y_{r,t}}{\partial k_t^2} = 0 && \frac{\partial^2 Y_{r,t}}{\partial k_t \partial A_t} = r_t L_t && \frac{\partial^2 Y_{r,t}}{\partial k_t \partial L_t} = r_t A_t \\
&\frac{\partial Y_{r,t}}{\partial A_t} = r_t k_t L_t && \frac{\partial^2 Y_{r,t}}{\partial A_t^2} = 0 && \frac{\partial^2 Y_{r,t}}{\partial A_t \partial L_t} = r_t k_t && \\
&\frac{\partial Y_{r,t}}{\partial L_t} = r_t k_t A_t && \frac{\partial^2 Y_{r,t}}{\partial L_t^2} = 0 && \frac{\partial^2 Y_{r,t}}{\partial r_t \partial L_t} = k_t A_t. && \qquad \text{(D.2.17)}
\end{aligned}$$

Using these and the results of stochastic calculus, i.e.

$$\begin{aligned}
dr_t dk_t &= (\theta-1)k_t r_t \sigma_{A+L}^2 dt & dr_t dA_t &= -(\theta-1)A_t r_t(\sigma_A^2 + \sigma_{AL})dt \\
dr_t dL_t &= -(\theta-1)L_t r_t(\sigma_L^2 + \sigma_{AL})dt & dA_t dk_t &= -A_t k_t(\sigma_A^2 + \sigma_{AL})dt \\
dL_t dk_t &= -L_t k_t(\sigma_L^2 + \sigma_{AL})dt & dA_t dL_t &= A_t L_t \sigma_{AL} dt,
\end{aligned}$$

in (D.2.16) gives

$$
\begin{aligned}
dY_{r,t} &= k_t A_t L_t (\theta - 1) r_t \\
&\quad \times \left\{ \left[s k_t^{\theta-1} - \alpha - n + \frac{\theta}{2}\sigma_{A+L}^2 - \sigma_{AL} \right] dt - (\sigma_A dz_A + \sigma_L dz_L) \right\} \\
&\quad + r_t A_t L_t k_t \left\{ \left[s k_t^{\theta-1} - (\alpha + n - \sigma_{A+L}^2 + \sigma_{AL}) \right] dt - (\sigma_A dz_A + \sigma_L dz_L) \right\} \\
&\quad + r_t k_t L_t (\alpha A_t dt + \sigma_A A_t dz_A) + r_t k_t A_t (n L_t dt + \sigma_L L_t dz_L) \\
&\quad + A_t L_t (\theta - 1) k_t r_t \sigma_{A+L}^2 dt - k_t L_t (\theta - 1) A_t r_t (\sigma_A^2 + \sigma_{AL}) dt \\
&\quad - k_t A_t (\theta - 1) L_t r_t (\sigma_L^2 + \sigma_{AL}) dt - r_t L_t A_t k_t (\sigma_A^2 + \sigma_{AL}) dt \\
&\quad - r_t A_t L_t k_t (\sigma_L^2 + \sigma_{AL}) dt + r_t k_t A_t L_t \sigma_{AL} dt
\end{aligned}
$$

and after some simplifications

$$
\begin{aligned}
&= Y_{r,t} \left[\theta s k_t^{\theta-1} + (1 - \theta) \left(\alpha + n - \frac{\theta}{2}\sigma_{A+L}^2 + \sigma_{AL} \right) \right] dt \\
&\quad + (1 - \theta) Y_{r,t} (\sigma_A dz_A + \sigma_L dz_L). \qquad \text{(D.2.18)}
\end{aligned}
$$

This is equation (5.3.32) in the text. Similarly, for the income from labor, $Y_{\omega,t} = \omega_t L_t$, Itô's lemma implies

$$
\begin{aligned}
dY_{\omega,t} &= \frac{\partial Y_{\omega,t}}{\partial \omega_t} d\omega_t + \frac{\partial Y_{\omega,t}}{\partial L_t} dL_t + \frac{1}{2} \left[\frac{\partial^2 Y_{\omega,t}}{\partial \omega_t^2} (d\omega_t)^2 + \frac{\partial^2 Y_{\omega,t}}{\partial L_t^2} (dL_t)^2 \right] \\
&\quad + \frac{\partial^2 Y_{\omega,t}}{\partial \omega_t \partial L_t} d\omega_t dL_t. \qquad \text{(D.2.19)}
\end{aligned}
$$

Using $d\omega_t dL_t = \omega_t L_t [(1 - \theta)\sigma_{AL} - \theta \sigma_L^2] dt + o(dt)$ as well as

$$
\frac{\partial Y_{\omega,t}}{\partial \omega_t} = L_t \quad \frac{\partial^2 Y_{\omega,t}}{\partial \omega_t^2} = 0 \quad \frac{\partial^2 Y_{\omega,t}}{\partial \omega_t \partial L_t} = 1 \quad \frac{\partial Y_{\omega,t}}{\partial L_t} = \omega_t \quad \frac{\partial^2 Y_{\omega,t}}{\partial L_t^2} = 0 \qquad \text{(D.2.20)}
$$

leads to

$$
\begin{aligned}
dY_{\omega,t} &= \omega_t L_t \left\{ \alpha dt + \theta \left[s k_t^{\theta-1} - \alpha - n - \frac{1}{2}(1 - \theta)\sigma_{A+L}^2 + \sigma_L^2 \right] dt \right. \\
&\quad \left. + [(1 - \theta)\sigma_A dz_A - \theta \sigma_L dz_L] \right\} \\
&\quad + \omega_t (n L_t dt + \sigma_L L_t dz_L) + \omega_t L_t [(1 - \theta)\sigma_{AL} - \theta \sigma_L^2] dt
\end{aligned}
$$

and finally to

$$
\begin{aligned}
&= Y_{\omega,t} \left[\theta s k_t^{\theta-1} + (1 - \theta) \left(\alpha + n - \frac{\theta}{2}\sigma_{A+L}^2 + \sigma_{AL} \right) \right] dt \\
&\quad + (1 - \theta) Y_{\omega,t} (\sigma_A dz_A + \sigma_L dz_L), \qquad \text{(D.2.21)}
\end{aligned}
$$

which is (5.3.33) in the main text.

Part V

Lists and References

List of Tables

List of Figures

Symbols and Abbreviations

Symbols introduced in Chap. 2

$A(t)$	Total factor productivity
$c(t)$	Consumption
$f(\circ)$	Intensive form of production function
$F(\circ)$	Aggregate production function
i	Index variable
k^*	Equilibrium value of $k(t)$
$k(t)$	Physical capital per capita
$K(t)$	Physical capital
L_0	Initial population size
$L(t)$	Human labor
$\tilde{M}(t+1)$	Stochastic discount factor
n	Population growth rate
$p_A(t)$	Price of asset A
$\tilde{R}_A(t+1)$	Return on asset A
$\tilde{R}_T(t+1)$	Return on total wealth
$\tilde{R}_W(t+1)$	Return on tradable wealth
s	Saving rate
t	Time (period)
$u_t(\circ)$	Instantaneous utility
$U(t)$	Utility function
$w_a(t)$	Wealth of a-aged agent
$\tilde{X}_A(t+1)$	Gross payoff of asset A
$y(t)$	Output (income) per capita
$y_T(t)$	Total income
$y_\omega(t)$	Labor income
$Y(t)$	Macroeconomic output
β	Subjective time discount factor
λ	Positive constant
θ	Cobb-Douglas Parameter

Symbols introduced in Chap. 3

a	Age index
$a_j(t)$	Age of generation j
$\tilde{A}(t)$	Random total factor productivity in consumption technology
B	Total factor productivity for capital adjustment
$c_a(t)$	Consumption of agent in generation a
$C_a(t)$	Aggregate consumption of generation a
$C(t)$	Aggregate consumption of all generations
$D(t)$	Dividends
$\mathfrak{g}(\circ)$	Actual investment in real capital net of adjustment costs
$\mathfrak{h}(\circ)$	Actual investment in parenthood net of adjustment costs
i	Index variable
$I(t)$	Gross Investment
$\mathfrak{I}(t)$	Information set
j	Index variable
$J_a[\circ]$	Indirect utility of a-aged agent
K_0	Initial physical capital stock
$K(t)$	Stock of physical capital
$\check{K}(t)$	Reduced form of equilibrium stock of physical capital
$L(t)$	Amount of human labor
$L^D(t)$	Demand for human labor
$L^S(t)$	Supply of human labor
$\tilde{M}(t+1)$	SDF of a-aged agent
$N_a(t)$	Number of agents in generation a
$\mathbb{N}$	Set of natural numbers
$p_H(t)$	Price of human capital security
$p_K(t)$	Price of real capital security
$p_Y(t)$	Price of consumable output
P_i	Probability i
$R(t)$	Return on riskfree asset
$\tilde{R}_H(t+1)$	Return on human capital security
$\tilde{R}_K(t+1)$	Return on real capital security
$\tilde{R}_M(t+1)$	Return on capital market
$\tilde{R}_{P,a}(t+1)$	Return on portfolio of a-aged agent
$s_a(t)$	Savings of a-aged agent
$S_a(t)$	Aggregated savings of generation a
$S(t)$	Aggregated savings of all generations
$S_H(t)$	Total savings in human capital
$S_K(t)$	Total savings in real capital
T	Life expectancy
$U_a(t)$	Lifetime utility function of a-aged agent
$u_a(\circ)$	Period utility function of a-aged agent
$w_a(t)$	Wealth of a-aged agent
$W_a(t)$	Aggregated wealth of generation a
$W(t)$	National wealth
$Y(t)$	Output of consumable units
$\check{Y}(t)$	Reduced form of equilibrium output
$Y^{GDP}(t)$	Gross domestic product
$\alpha_{H,a}(t)$	Portfolio allocation to human capital by a-aged agent
$\alpha_{K,a}(t)$	Portfolio allocation to real capital by a-aged agent
$\alpha_H^M(t)$	Weight of human capital securities in market portfolio
$\alpha_K^M(t)$	Weight of real capital securities in market portfolio
β	Subjective time discount factor
γ	Coefficient of relative risk aversion
δ_H	Illustrative "depreciation" rate in labor population
δ_K	Illustrative depreciation rate of real capital
$\tilde{\theta}(t)$	Random elasticity parameter in consumption technology
ϑ	Elasticity parameter for capital adjustment
Ξ_i	Auxiliary constant i
$\Pi(t)$	Profits of firms
$\psi(t)$	Investment-output ratio
Ψ	Reduced form of equilibrium investment-output ratio
$\omega(t)$	Wage rate

Symbols introduced in Chap. 4

Prime notation characterizes same variables as above for mimicking economy.

$i'_a(t)$	Generation a's fraction of physical capital investments	$\mathfrak{t}_y(t)$	Transfer benefits of young
$l(t)$	Fraction of work period spent in leisure	$\mathfrak{T}_o(t)$	Budget of PAYGO system
$\tilde{R}'_{T,a}(t+1)$	Return on artificial transfer portfolio of a-aged agent	$\mathfrak{T}_y(t)$	Budget of fiscal system
$\tilde{\mathfrak{R}}(t+1)$	Implicit return on PAYGO	Γ	Scaling factor for utils from leisure
$\mathfrak{t}_o(t)$	PAYGO benefits of old	τ_C	Consumption tax rate
		τ_ω	PAYGO contribution rate

Symbols introduced in Chap. 5

A_0	Initial technology level	Y_t	Instantaneous consumable output
A_t	Instantaneous technology input	$Y_{r,t}$	Total factor income of real capital
C_t	Instantaneous consumption	$Y_{\omega,t}$	Total factor income of human labor
dz_i	Standard Wiener process i		
$G'(t)$	Illustrative government expenditure		
I_t	Instantaneous investments		
k_t	Stock of physical capital per capita	α	Drift parameter of technology process
K_t	Stock of physical capital	δ	Instantaneous depreciation
L_0	Initial level of human labor	Π_t	Instantaneous profits of firms
L_t	Instantaneous input of human labor	ρ_{z_L,z_A}	Instantaneous correlation of dz_A and dz_L
n	Drift parameter of process for human labor	σ_A	Volatility of technology process
$p_{Y,t}$	Instantaneous price of consumption units	σ_{AL}	Instantaneous covariance of dz_A and dz_L
r_t	Instantaneous rental of physical capital	σ^2_{A+L}	Variance of $dz_A + dz_L$
$\tilde{R}_{H,t}$	Return on human capital security	σ_L	Volatility of labor process
$\tilde{R}_{K,t}$	Return on physical capital security	ω_t	Instantaneous wage rate

Other Abbreviations

etc.	Et cetera	n.m.	Not meaningful
i.e.	That is	PAYGO	Pay-as-you-go
n.a.	Not applicable	SDF	Stochastic discount factor

References

Abel, A. B. [1983], 'Optimal investment under uncertainty', *American Economic Review* **73**(1), 228–233.

Abel, A. B. [2001], 'Will bequest attenuate the predicted meltdown in stock prices when baby boomers retire?', *Review of Economics and Statistics* **83**(4), 589–595.

Abel, A. B. [2003], 'The effects of the baby boom on stock prices and capital accumulation in the presence of Social Security', *Econometrica* **71**(2), 551–578.

Abel, A. B. and Blanchard, O. J. [1986], 'The present value of profits and cyclical movements in investment', *Econometrica* **54**(2), 249–273.

Abel, A. B. and Eberly, J. C. [1994], 'A unified model of investment under uncertainty', *American Economic Review* **84**(5), 1369–1384.

Abel, A. B. and Eberly, J. C. [1996], 'Optimal investment with costly reversibility', *Review of Economic Studies* **63**(4), 581–593.

Aghion, P. and Howitt, P. [1998], *Endogenous Growth Theory*, The MIT Press, Cambridge MA (USA).

Allais, M. [1947], *Economy and Interest*, Imprimerie Nationale et Librairie des Publications Officielles, Paris (France). "Économie et Intérêt", two volumes.

Alvarado, J. and Creedy, J. [1998], *Population Ageing, Migration and Social Expenditure*, Edward Elgar, Northampton MA (USA).

Ameriks, J. and Zeldes, S. P. [2001], How do household portfolio shares vary with age? Working Paper 6-120101, TIAA-CREF Institute.

Ang, A. and Maddaloni, A. [2003], Do demographic changes affect risk premiums? Evidence from international data. NBER Working Paper No. 9677, National Bureau of Economic Research.

Arrow, K. [1971], Aspects of the theory of risk-bearing, *in* K. Arrow, ed., 'Essays in the Theory of Risk-Bearing', Markham, Chicago IL (USA). based on Yrjo Hahnsson Foundation lectures [1965].

Arrow, K. and Debreu, G. [1954], 'Existence of an equilibrium for a competitive economy', *Econometrica* **22**(3), 265–290.

Auerbach, A. J., Gokhale, J. and Kotlikoff, L. J. [1991], Generational accounts: A meaningful alternative to deficit accounting, *in* D. Bradford, ed., 'Tax Policy and the Economy:', Vol. 5 of *NBER Tax Policy and the Economy*, The MIT Press, Cambridge MA (USA), pp. 55–110.

Auerbach, A. J., Gokhale, J. and Kotlikoff, L. J. [1992], 'Generational accounting: A new approach for understanding the effects of fiscal policy on saving', *Scandinavian Journal of Economics* **94**(2), 303–318.

Auerbach, A. J., Gokhale, J. and Kotlikoff, L. J. [1994], 'Generational accounting: A meaningful way to evaluate fiscal policy', *Journal of Economic Perspectives* **8**(1), 73–94.

Auerbach, A. J. and Herrmann, H., eds [2002], *Ageing, financial markets and monetary policy*, Springer, Berlin (Germany).

Auerbach, A. J. and Kotlikoff, L. J. [1987], *Dynamic Fiscal Policy*, Cambridge University Press, Cambridge (UK).

Auerbach, A. J., Kotlikoff, L. J. and Leibfritz, W., eds [1999], *Generational Accounting around the World*, NBER Project Report, University of Chicago Press, Chicago IL (USA).

Bakshi, G. S. and Chen, Z. [1994], 'Baby boom, population aging, and capital markets', *Journal of Business* **67**(2), 165–202.

Balasko, Y. and Shell, K. [1980], 'The overlapping-generations model i: The case of pure exchange without money', *Journal of Economic Theory* **23**(3), 281–306.

Balasko, Y. and Shell, K. [1981*a*], 'The overlapping-generations model ii: The case of pure exchange with money', *Journal of Economic Theory* **24**(1), 112–142.

Balasko, Y. and Shell, K. [1981*b*], 'The overlapping-generations model iii: The case of log-linear utility functions', *Journal of Economic Theory* **24**(1), 143–152.

Balvers, R. J., Cosimano, T. F. and McDonald, B. [1990], 'Predicting stock return in an efficient market', *Journal of Finance* **45**(4), 1109–1128.

Barberis, N. [2000], 'Investing for the long run when returns are predictable', *Journal of Finance* **55**(1), 225–264.

Barro, R. J. and Becker, G. S. [1989], 'Fertility choice in a model of economic growth', *Econometrica* **57**(2), 481–501.

Barro, R. J. and Sala-i Martin, X. [2004], *Economic Growth*, 2nd edn, The MIT Press, Cambridge MA (USA).

Basu, P. [1987], 'An adjustment cost model of asset pricing', *International Economic Review* **28**(3), 609–621.

Baxter, M. and King, R. G. [2001], The role of international investments in a privatized Social Security system, *in* J. Y. Campbell and M. Feldstein, eds, 'Risk-Aspects of Investment-based Social Security Reform', NBER Books, University of Chicago Press, Chicago IL (USA), pp. 371–437.

Becker, G. S. [1960], An economic analysis of fertility, *in* A. J. Coale, ed., 'Demographic and Economic Change in Developed Countries', Vol. 11 of

NBER Conference Series, Princeton University Press, Princeton NJ (USA), pp. 209–231.

Becker, G. S. [1962], 'Investment in human capital: A theoretical analysis', *Journal of Political Economy* **50**(5-2), 9–49.

Becker, G. S. [1991], *A Treatise on the Family*, Harvard University Press, Cambridge MA (USA).

Becker, G. S. [1993], *Human Capital: A Theoretical and Empirical Analysis, with Special Reference to Education*, 3rd edn, University of Chicago Press, Chicago IL (USA).

Becker, G. S. and Lewi, G. H. [1973], 'On the interaction between the quantity and quality of children', *Journal of Political economy* **81**(2), 279–288.

Becker, G. S., Murphy, K. M. and Tamura, R. [1990], 'Human capital, fertility, and economic growth', *Journal of Political Economy* **98**(5), 12–37.

Becker, G. S. and Tomes, N. [1986], 'Human capital and the rise and fall of families', *Journal of Labor Economics* **4**(3-2), 1–39.

Bergantino, S. M. [1998], Life Cycle Investment Behavior, Demographics and Asset Prices, PhD thesis, Massachusetts Institute of Technology, Cambridege MA (USA).

Bernoulli, D. [1738], *Commentarii Academiae Scientiarum Imperialis Petropolitanae*, Vol. 5, Imperial Academy of Sciences in St. Petersburg, St. Petersburg (Russia), chapter Specimen Theoriae Novae de Mensura Sortis, pp. 175–192. Translated into English by Louise Sommer [1954] "Exposition of a New Theory on the Measurement of Risk," *Econometrica* **22**(1), 23-3.

Birg, H. [2003], *Die demographische Zeitenwende - Der Bevölkerungsrückgang in Deutschland und Europa*, 3rd edn, C.H. Beck, Munich (Germany).

Black, F. [1972], 'Capital market equilibrium with restricted borrowing', *Journal of Business* **45**(3), 444–455.

Black, F. and Scholes, M. S. [1973], 'The pricing of options and corporate liabilities', *Journal of Political Economy* **81**(3), 637–654.

Blacker, C. P. [1947], 'Stages in population growth', *Eugenics Review* **39**(3), 88–101.

Blanchard, O. J. [1985], 'Debt, deficit, and finite horizons', *Journal of Political Economy* **93**(2), 223–247.

Bodie, Z., Merton, R. C. and Samuelson, W. F. [1992], 'Labor supply flexibility and portfolio choice in a life cycle model', *Journal of Economic Dynamics and Control* **16**(3-4), 427–449.

Bohn, H. [1998], Risk sharing in a stochastic overlapping generations economy. Working Paper, University of California at Santa Barbara.

Boldrin, M., Christiano, L. and Fisher, J. [2001], 'Habit persistence, asset returns, and the business cycle', *American Economic Review* **91**(1), 149–166.

Borgmann, C. [2005], *Social Security, Demographics, and Risk*, Population Economics, Springer, Berlin (Germany).

Bourguignon, F. [1974], 'A particular class of continuous-time stochastic growth models', *Journal of Economic Theory* **9**(2), 141–158.

Brainard, W. and Tobin, J. [1968], 'Pitfalls in financial model building', *American Economic Review* **58**(2), 99–122.

Breeden, D. T. [1979], 'An intertemporal asset pricing model with stochastic consumption and investment opportunities', *Journal of Financial Economics* **7**(3), 265–296.

Breeden, D. T. [1986], 'Consumption, production, inflation, and interest rates: A synthesis', *Journal of Political Economy* **16**(1), 3–39.

Brock, W. A. [1979], An integration of stochastic growth theory and theory of finance: Part i: The growth model, *in* J. R. Green and J. A. Scheinkman, eds, 'General Equilibrium, Growth, and Trade - Essays in Honor of Lionel McKenzie', Economic Theory, Econometrics, and Mathematical Economics, Academic Press, New York NY (USA), pp. 165–192.

Brock, W. A. [1982], Asset prices in a production economy, *in* J. J. McCall, ed., 'The Economics of Information and Uncertainty - A Conference Report', Vol. 2 of *NBER Conference Series*, University of Chicago Press, Chicago IL (USA), pp. 1–46.

Brock, W. A. and Mirman, L. J. [1972], 'Optimal economic growth and uncertainty: The discounted case', *Journal of Economic Theory* **4**(3), 479–513.

Brooks, R. [2000], What will happen to financial markets when the baby boomers retire? IMF Working Paper No. 00/18, International Monetary Fund.

Brooks, R. [2002], 'Asset-market effects of the baby boom and Social Security', *American Economic Review* **92**(2), 402–406.

Brooks, R. [2003], 'Population aging and global capital flows in a parallel universe', *IMF Staff Papers* **50**(3), 200–221. International Monetary Fund.

Brown, L. R. [1992], *Who will feed China? Wake-up call for a small planet*, The Worldwatch Environmental Alert Series, W.W. Norton & Company, New York NY (USA).

Börsch-Supan, A. [2001], Rentabilitätsvergleiche im Umlage- und Kapitaldeckungsverfahren: Konzepte, empirische Ergebnisse, sozialpolitische Konsequenzen, *in* E. Theurl, ed., 'Sozialstaat an der Jahrtausendwende - Analysen und Perspektiven', Physica, Heidelberg (Germany), pp. 207–233.

Börsch-Supan, A. [2003], 'Labor market effects of population aging', *Review of Labor Economics and Industrial Relations* **17**(Special Issue), 5–44.

Börsch-Supan, A. H. and Wilke, C. B. [2003], The german public pension system: How it was, how it will be. Working Paper No. 34-2003, Mannheim Research Institute for the Economics of Aging (MEA), University of Mannheim.

Börsch-Supan, A., Heiss, F., Ludwig, A. and Winter, J. [2002], Ageing and international capital flows, *in* A. J. Auerbach and H. Herrmann, eds, 'Ageing, financial markets and monetary policy', Springer, Berlin (Germany), pp. 55–83.

Börsch-Supan, A., Heiss, F., Ludwig, A. and Winter, J. [2003], 'Pension reform, capital markets and the rate of return', *German Economic Review* **4**(2), 151–181.

Börsch-Supan, A., Ludwig, A. and Winter, J. [2005], 'Aging, pension reform, and capital flows: A multi-country simulation model', *Economica* . Forthcoming.

Bubb, A. and Zimmermann, H. [2002], 'Das Risiko der Vorsorge (Lange integrale Fassung)', Avenir Suisse, Zurich (Switzerland).

Bundesministerium für Gesundheit und Soziale Sicherung, ed. [2003], *Nachhaltigkeit in der Finanzierung der Sozialen Sicherungssysteme - Bericht der Komission*, Bundesministerium für Gesundheit und Soziale Sicherung, Berlin (Germany).

Campbell, J. Y. [1993], 'Intertemporal asset pricing without consumption data', *American Economic Review* **83**(3), 487–512.

Campbell, J. Y. [1994], 'Inspecting the mechanism: an analytical approach to the stochastic growth model', *Journal of Monetary Economics* **33**(3), 463–506.

Campbell, J. Y. [1996], 'Understanding risk and return', *Journal of Political Economy* **104**(2), 298–345.

Campbell, J. Y. [2000], 'Asset pricing at the millennium', *Journal of Finance* **55**(5), 1515–1567.

Campbell, J. Y. [2001], 'A comment on James M. Porterba's "Demographic structure and asset returns"', *Review of Economics and Statistics* **83**(4), 586–588.

Campbell, J. Y. [2003*a*], Consumption-based asset pricing, *in* G. M. Constantinides, M. Harris and R. M. Stulz, eds, 'Handbook of the Economics of Finance', Vol. 1B - Financial Markets and Asset Pricing of *Handbooks in Economics*, Elsevier, North Holland, Amsterdam (Netherlands), pp. 803–887.

Campbell, J. Y. [2003*b*], The equity premium puzzle in retrospect, *in* G. M. Constantinides, M. Harris and R. M. Stulz, eds, 'Handbook of the Economics of Finance', Vol. 1B - Financial Markets and Asset Pricing of *Handbooks in Economics*, Elsevier, North Holland, Amsterdam (Netherlands), pp. 888–936.

Campbell, J. Y., Chacko, G., Rodriguez, J. and Viceira, L. M. [2004], 'Strategic asset allocation in a continuous-time var model', *Journal of Economic Dynamics and Control* **28**(11), 2195–2214.

Campbell, J. Y., Chan, Y. and Viceira, L. M. [2003], 'A multivariate model of strategic asset allocation', *Journal of Financial Economics* **67**(1), 41–80.

Campbell, J. Y., Cocco, J. F., Gomes, F. J. and Maenhout, P. J. [2001], Investing retirement wealth - a life-cycle model, *in* J. Y. Campbell and M. Feldstein, eds, 'Risk Aspects of Investment-based Social Security Reform', NBER Project Report, University of Chicago Press, Chicago IL (USA), pp. 439–473.

Campbell, J. Y., Cocco, J. F., Gomes, F. J. and Maenhout, P. J. [2005], 'Consumption and portfolio choice over the life-cycle', *Review of Financial Studies* . Forthcoming.

Campbell, J. Y. and Feldstein, M., eds [2004], *Risk Aspects of Investment-based Social Security Reform*, NBER Project Report, University of Chicago Press, Chicago IL (USA).

Campbell, J. Y., Lo, A. W. and MacKinlay, A. C. [1997], *The Econometrics of Financial Markets*, Princeton University Press, Princeton NJ (USA).

Campbell, J. Y. and Viceira, L. M. [1999], 'Consumption and portfolio decisions when expected returns are time varying', *Quarterly Journal of Economics* **114**(2), 433–495.

Campbell, J. Y. and Viceira, L. M. [2001*a*], *Strategic Asset Allocation. Portfolio Choice for long-term Investors*, Clarendon Lectures in Economics, Oxford University Press, Oxford (UK).

Campbell, J. Y. and Viceira, L. M. [2001*b*], 'Who should buy long-term bonds?', *American Economic Review* **91**(1), 99–127.

Campbell, J. Y., Viceira, L. M. and White, J. [2003], 'Foreign currency for long-term investors', *Economic Journal* **113**(486), C1–C25.

Cass, D. [1965], 'Optimum growth in an aggregative model of capital accumulation', *Review of Economic Studies* **32**(3), 233–240.

Cass, D. and Shell, K. [1980], In defense of a basic approach, *in* J. Kareken and N. Wallace, eds, 'Models of Monetary Economies', Federal Reserve Bank of Minneapolis, Minneapolis MN (USA), pp. 251–260.

Cass, D. and Stiglitz, J. E. [1970], 'The structure of investor preferences and asset returns, and separability in portfolio allocation: Contribution to the pure theory of mutual funds', *Journal of Economic Theory* **2**(2), 122–160.

Chand, S. K. and Jaeger, A. [1996], Aging populations and public pension schemes. IMF Occasional Paper No. 147, International Monetary Fund.

Chang, F.-R. and Malliaris, A. G. [1987], 'Asymptotic growth under uncertainty: Existence and uniqueness', *Review of Economic Studies* **54**(1), 169–174.

Cheng, K. C. [2003], Economic implications of china's demographics in the 21st century. IMF Working Paper No. 03/29, International Monetary Fund.

Clark, J. B. [1899], *The Distribution of Wealth: A Theory of Wages, Interest and Profits*, 1908 edn, Macmillan, New York NY (USA).

Cocco, J. F. [2005], 'Portfolio choice in the presence of housing', *Review of Financial Studies* . Forthcoming.

Cochrane, J. H. [1991], 'Production-based asset pricing and the link between stock returns and economic fluctuations', *Journal of Finance* **46**(1), 209–237.

Cochrane, J. H. [1996], 'A cross-sectional test of an investment-based asset pricing model', *Journal of Political Economy* **104**(3), 572–621.

Cochrane, J. H. [2001], *Asset Pricing*, Princeton University Press, Princeton NJ (USA).

Cochrane, J. H. [2005], Financial markets and the real economy. NBER Working Paper No. 11193, National Bureau of Economic Research.

Cochrane, J. H. and Hansen, L. P. [1992], Asset pricing lessons for macroeconomics, *in* O. J. Blanchard and S. Fischer, eds, 'NBER Macroeconomics

Annual 1992', National Bureau of Economic Research, The MIT Press, Cambridge MA (USA), pp. 115–169.

Constantinides, G. M. [1980], 'Admissible uncertainty in the capital asset pricing model', *Journal of Financial Economics* **8**(1), 71–86.

Constantinides, G. M. [1982], 'Intertemporal asset pricing with heterogeneous consumers and without demand aggregation', *Journal of Business* **55**(2), 253–267.

Constantinides, G. M. [1989], Theory of valuation: Overview and recent developments, *in* S. Bhattacharya and G. M. Constantinides, eds, 'Theory of Valuation: Frontiers of Modern Financial Theory', Vol. 1 of *Rowman and Littlefiels Studies in Financial Economics*, Rowman and Littlefiels, Totowa NJ (USA), chapter 1, pp. 1–23.

Constantinides, G. M. [1990], 'Habit formation: A resolution of the equity premium puzzle', *Journal of Political Economy* **98**(3), 519–543.

Constantinides, G. M. [1992], 'A theory of the nominal term structure of interest rates', *Review of Financial Studies* **5**(4), 531–552.

Constantinides, G. M. [2002], 'Rational asset prices', *Journal of Finance* **57**(4), 1567–1591.

Constantinides, G. M., Donaldson, J. B. and Mehra, R. [2002], 'Junior can't borrow: A new perspective on the equity premium puzzle', *Quarterly Journal of Economics* **117**(1), 269–296.

Constantinides, G. M., Donaldson, J. B. and Mehra, R. [2005*a*], Junior is rich: Bequests as consumption. NBER Working Paper No. 11122, National Bureau of Economic Research.

Constantinides, G. M., Donaldson, J. B. and Mehra, R. [2005*b*], 'Junior must pay: Pricing the implicit put in privatizing Social Security', *Annals of Finance* **1**(1), 1–34.

Constantinides, G. M. and Malliaris, A. G. [1995], Portfolio theory, *in* R. A. Jarrow, W. Maksimovic and W. T. Ziemba, eds, 'Finance', Vol. 9 of *Handbooks in Operations Research and Management Science*, Elsevier, North Holland, Amsterdam (Netherlands), pp. 1–30.

Copeland, T. E. and Weston, J. F. [1988], *Financial Theory and Corporate Policy*, reprinted 1992 edn, Addison-Wesley.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. [1985*a*], 'An intertemporal general equilibrium model of asset prices', *Econometrica* **53**(2), 363–384.

Cox, J. C., Ingersoll, J. E. and Ross, S. A. [1985*b*], 'A theory of the term structure of interest rates', *Econometrica* **53**(2), 385–408.

Cox, J. C. and Ross, S. A. [1976], 'The valuation of options for alternative stochastic processes', *Journal of Financial Economics* **3**(1), 145–166.

Cragg, R. [1998], *The Demographic Investor – Strategies for Surviving the Pension Crises*, Financial Times Pitman, London (UK).

Cullis, J. and Jones, P. [1998], *Public Finance and Public Choice*, 2nd edn, Oxford University Press, Oxford (UK).

Currais, L. [2000], 'From the malthusian regime to the demographic transition: Contemporary research and beyond', *Econômica* **2**(3), 75–103.

Davis, E. P. [2004], Demographic and pension-system challenges to monetary and financial stability. Working paper presented at the Austrian National Bank economics conference, May 2004.

Davis, E. P. and Li, C. [2003], Demographics and financial asset prices in the major industrial economies. Finance Discussion Paper No. 03/07, Brunel University.

Day, J. C. [1996], *Population Projections of the United States by Age, Sex, Race, and Hispanic Origin: 1995 to 2050*, Current Population Reports, No. P25-1130, U.S. Census Bureau, Government Printing Office, Washington D.C. (USA).

Demange, G. and Laroque, G. [1999], 'Social Security and demographic shocks', *Econometrica* **67**(3), 527–542.

Diamond, P. A. [1965], 'National debt in a neoclassical growth model', *American Economic Review* **55**(5), 1126–1150.

Diamond, P. A. [2000], 'What stock market returns to expect for the future?', *Social Security Bulletin* **63**(2), 38–52.

Dixit, A. K. and Pindyck, R. S. [1994], *Investment under Uncertainty*, Princeton University Press, Princeton NJ (USA).

Domar, E. D. [1946], 'Capital expansion, rate of growth, and unemployment', *Econometrica* **14**, 137–147.

Domeij, D. and Floden, M. [2003], Population aging and international capital flows. Working Paper, Stockholm School of Economics.

Domeij, D. and Floden, M. [2005], More on population aging and international capital flows. Working Paper, Stockholm School of Economics.

Donaldson, P. J. [1999], *World Population: More than just numbers*, Population Reference Bureau, Washington D.C. (USA).

Duesenberry, J. S. [1949], *Income, Saving and the Theory of Consumer Behavior*, Harvard Economic Studies, Harvard University Press, Cambridge MA (USA).

Duffie, D. [2001], *Dynamic Asset Pricing Theory*, 3rd edn, Princeton University Press, Princeton NJ (USA).

Dunaway, S. V. and N'Diaye, P. M. [2004], An approach to long-term fiscal policy analysis. IMF Working Paper No. 04/113, International Monetary Fund.

Dutta, J., Kapur, S. and Orszag, M. J. [2000], 'A portfolio approach to the optimal funding of pensions', *Economics Letters* **69**(2), 201–206.

Easterlin, R. A. [1968], *Population, labor force, and long swings in economic growth: the American experience*, National Bureau of Economic Research, Columbia University Press, New York NY (USA).

Eisner, R. and Strotz, R. H. [1963], Determinants of investment, *in* Commission on Money and Credit, ed., 'Impacts of Monetary Policy', Prentice-Hall, Englewood Cliffs NJ (USA), pp. 59–337.

England, R. S. [2002*a*], *Global Aging and Financial Markets - Hard Landings ahead?*, Center for Strategic and International Studies, CSIS Press, Washington D.C. (USA).

England, R. S. [2002*b*], *The Macroeconomic Impact of Global Aging - A new Era of Economic Frailty?*, Center for Strategic and International Studies, CSIS Press, Washington D.C. (USA).

England, R. S. [2005], *Aging China*, Center for Strategic and International Studies, CSIS Press, Washington D.C. (USA).

European Commission [1999], *Generational accounting in Europe*, Economic Policy Committee and the Directorate-General for Economic and Financial Affairs of the European Commission, Brussels (Belgium). European Economy, Reports and Studies, No. 6/1999.

European Commission [2001], *Budgetary challenges posed by ageing populations: the impact on public spending on pensions, health and long-term care for the elderly and possible indicators of the long-term sustainability of public finances*, Economic Policy Committee and the Directorate-General for Economic and Financial Affairs of the European Commission, Brussels (Belgium). European Economy, Reports and Studies, No. 4/2001.

Fama, E. F. [1970*a*], 'Efficient capital markets: A review of theory and empirical work', *Journal of Finance* **25**(2), 383–417.

Fama, E. F. [1970*b*], 'Multiperiod consumption-investment decisions', *American Economic Review* **60**(1), 163–174.

Fama, E. F. [1976], *Foundations of Finance: Portfolio Decision and Securities Prices*, Basic Books, New York NY (USA).

Fama, E. F. and French, K. R. [1989], 'Business conditions and expected returns on stocks and bonds', *Journal of Financial Economics* **25**(1), 23–49.

Fama, E. F. and French, K. R. [1992], 'The cross-section of stock returns', *Journal of Finance* **47**(2), 427—465.

Fama, E. F. and French, K. R. [1993], 'Common risk factors in stock and bond returns', *Journal of Financial Economics* **33**(1), 3–56.

Fama, E. F. and French, K. R. [1995], 'Size and book-to-market factors in earnings and returns', *Journal of Finance* **50**(1), 131–155.

Fama, E. F. and French, K. R. [1996], 'Multifactor explanations of asset pricing anomalies', *Journal of Finance* **51**(1), 55–84.

Fama, E. F. and Miller, M. H. [1972], *The Theory of Finance*, Dryden Press, Hinsdale IL (USA).

Fama, E. F. and Schwert, W. G. [1977], 'Human capital and capital market equilibrium', *Journal of Financial Economics* **4**(1), 95–125.

Fehr, H. and Habermann, C. [2004], Pension reform and demographic uncertainty: The case of Germany. Würzburg Economic Papers, No.47.

Feldstein, M. [1974], 'Social Security, induced retirement, and aggregate capital accumulation', *Journal of Political Economy* **82**(5), 905–926.

Feldstein, M. [1976], 'Social Security and saving: The extended life cycle theory', *American Economic Review* **66**(2), 77—86.

Feldstein, M. [1980], 'International differences in Social Security and saving', *Journal of Public Economics* **14**(2), 225–244.

Feldstein, M. [1985], 'The optimal level of Social Security benefits', *Quarterly Journal of Economics* **10**(2), 303–320.

Feldstein, M. [1987], 'Should Social Security be means tested?', *Journal of Political Economy* **95**(3), 468–484.

Feldstein, M., ed. [1988], *Privatizing Social Security*, NBER Project Report, University of Chicago Press, Chicago IL (USA).

Feldstein, M. and Horioka, C. [1980], 'Domestic savings and international capital flows', *Economic Journal* **90**(358), 314–329.

Feldstein, M. and Liebman, J. B., eds [2002], *The Distributional Aspects of Social Security and Social Security Reform*, NBER Project Report, University of Chicago Press, Chicago IL (USA).

Feldstein, M. and Siebert, H., eds [2002], *Social Security Pension Reform in Europe*, NBER Project Report, University of Chicago Press, Chicago IL (USA).

Feroli, M. [2003], Capital flows among the G-7 nations: A demographic perspective. Finance and Economics Discussion Series, Working Paper No. 03/54, Federal Reserve Board.

Fisher, I. [1930], *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It*, Macmillan, New York NY (USA).

Fitoussi, J.-P. [2003], Monetary policy, the promotion of growth and the SGP in an ageing society. Briefing paper No. 4 –14, Theme No. 2, European Parliament, Committee for economic and monetary affairs.

Fleming, W. H. and Rishel, R. W. [1975], *Deterministic and Stochastic Optimum Control*, Vol. 1 of *Applications of Mathematics*, Springer, New York NY (USA).

Fogel, R. W. [1994], 'Economic growth, population theory, and physiology: The bearing of long-term processes on the making of economic policy', *American Economic Review* **84**(3), 369–395.

Fogel, R. W. [1997], 'Economic and social structure for an ageing population', *Philosophical Transactions: Biological Sciences* **352**(1363), 1905–1917.

Fogel, R. W. [2003], 'Forecasting the demand for health care in OECD nations and China', *Contemporary Economic Policy* **21**(1), 1–10.

Fogel, R. W. and Costa, D. L. [1997], 'A theory of technophysio evolution, with some implications for forecasting population, health care costs, and pension costs', *Demography* **34**(1), 49–66.

Frankel, J. [1991], Quantifying international capital mobility in the 1980s, *in* B. D. Bernheim and J. B. Shoven, eds, 'National Saving and Economic Performance', NBER Project Report, University of Chicago Press, Chicago IL (USA), pp. 227–269.

French, K. R. and Poterba, J. M. [1991], 'Investor diversification and international equity markets', *American Economic Review* **81**(2), 222–226.

Friedman, M. [1957], *A Theory of the Consumption Function*, NBER General Series No. 63, Princeton University Press, Princeton NJ (USA).

Friedman, M. [1962], *Capitalism and Freedom*, 40th anniversary (2002) edn, University of Chicago Press, Chicago IL (USA).

Gale, D. [1973], 'Pure exchange equilibrium of dynamic economic models', *Journal of Economic Theory* **6**(1), 12–36.

Geanakoplos, J., Magill, M. and Quinzii, M. [2002], Demography and the long-run predictability of the stock market. Cowles Foundation Discussion Paper No. 1380, Cowles Foundation/Yale University.

Gollier, C. [2001], *The Economics of Risk and Time*, The MIT Press, Cambridge MA (USA).

Gomes, F. and Michaelides, A. [2004], A human capital explanation for an asset allocation puzzle. UBS-PRP Discussion Paper No. 24, London School of Economics.

Gomes, J. F., Kogan, L. and Zhang, L. [2003], 'Equilibrium cross-section of returns', *Journal of Political Economy* **111**(4), 693–732.

Gould, J. P. [1968], 'Adjustment costs in the theory of investment of the firm', *Review of Economic Studies* **35**(1), 47–55.

Goyal, A. [2002], 'Demographics, stock market flows, and stock returns', *Journal of Financial and Quantitative Analysis* **39**(1), 115–142.

Greenwood, J., Seshadri, A. and Vandenbroucke, G. [2005], 'The baby boom and baby bust: Some macroeconomics for population economics', *American Economic Review* . Forthcoming.

Grossman, S. J. and Shiller, R. J. [1981], 'The determinants of the variability of stock market prices', *American Economic Review* **71**(2), 222–227.

Group of Ten [1998], *The macroeconomic and financial implications of ageing populations*, Bank for International Settlements, Basel (Switzerland).

Hakansson, N. H. [1970], 'Optimal investment and consumption strategies under risk for a class of utility functions', *Econometrica* **38**(5), 587–607.

Hansen, L. P. and Jagannathan, R. [1991], 'Implications of security market data for models of dynamic economies', *Journal of Political Economy* **99**(2), 225–262.

Hansen, L. P. and Richard, S. F. [1981], 'The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models', *Econometrica* **55**(3), 587–613.

Harrison, J. M. and Kreps, D. M. [1979], 'Martingale and arbitrage in multiperiods securities markets', *Journal of Economic Theory* **20**(3), 381–408.

Harrod, R. F. [1939], 'An essay in dynamic theory', *Economic Journal* **49**, 14–33.

Harrod, R. F. [1948], *Towards a Dynamic Economics: Some recent Developments of Economic Theory and their Application to Policy*, Macmillan, London (UK).

Haupt, A. and Kane, T. T. [2004], *Population Reference Bureau's Population Handbook*, 5th edn, Population Reference Bureau, Washington D.C. (USA).

Hayashi, F. [1982], 'Tobin's average q and marginal q: A neoclassical interpretation', *Econometrica* **50**(1), 213–224.

Heaton, J. and Lucas, D. [1997], 'Market frictions, savings behavior, and portfolio choice', *Macroeconomic Dynamics* **1**(1), 76–101.

Heaton, J. and Lucas, D. [2000], 'Portfolio choice in the presence of background risk', *Economic Journal* **110**(460), 1–26.

Henry, J. F. [1990], *The making of neoclassical economics*, Unwin Hyman, Boston MA (USA).

Hicks, J. R. [1932], *The Theory of Wages*, Macmillan, London (UK).

Hicks, J. R. [1939], *Value and Capital: An inquiry into some fundamental principles of economic theory*, 2nd (1946) edn, Clarendon Press, Oxford (UK).

Hirshleifer, J. [1958], 'On the theory of optimal investment decision', *Journal of Political Economy* **66**(4), 329–352.

Hirshleifer, J. [1970], *Investment, Interest, and Capital*, Prentice-Hall, Englewood Cliffs NJ (USA).

Holzmann, R. [2000], 'The world bank approach to pension reform', *International Social Security Review* **53**(1), 11–34.

Holzmann, R. [2004], Toward a Reformed and Coordinated Pension System in Europe: Rationale and Potential Structure. Social Protection Discussion Paper Series No. 0407, The World Bank.

Holzmann, R. and Hinz, R. [2005], *Old-Age Income Support in the Twenty-first Century: An International Perspective on Pension Systems and Reform*, World Bank, Washington D.C. (USA). Joint with Hermann von Gersdorff, Indermit Gill, Gregorio Impavido, Alberto R. Musalem, Michal Rutkowski, Robert Palacios, Yvonne Sin, Kalanidhi Subbarao, Anita Schwarz.

Holzmann, R. and Stiglitz, J. E., eds [2001], *New Ideas about Old Age Security. Towards Sustainable Pension Systems in the 21st Century*, The World Bank, Washington D.C. (USA).

Inada, K.-I. [1963], 'On a two-sector model of economic growth: Comments and a generalization', *Review of Economic Studies* **30**(2), 119–127.

Ingersoll, J. E. [1987], *Theory of Financial Decision Making*, Studies in Financial Economics, Rowman & Littlefield, Savage MD (USA).

ISSA [2003*a*], *Social Security Programs Throughout the World: Africa, 2003*, International Social Security Association and Social Security Administration, Geneva (Switzerland).

ISSA [2003*b*], *Social Security Programs Throughout the World: Asia and the Pacific, 2002*, International Social Security Association and Social Security Administration, Geneva (Switzerland).

ISSA [2004*a*], *Social Security Programs Throughout the World: Europe, 2004*, International Social Security Association and Social Security Administration, Geneva (Switzerland).

ISSA [2004*b*], *Social Security Programs Throughout the World: The Americas, 2003*, International Social Security Association and Social Security Administration, Geneva (Switzerland).

Jackson, R. [2002], *The Global Retirement Crisis - The Threat to World Stability and What to Do About It*, The Center for Strategic and International Studies and CitiGroup, Washington D.C. (USA).

Jackson, R. [2003], *Germany and the Challenge of Global Aging*, The Center for Strategic and International Studies and Nationwide Global, Washington D.C. (USA).

Jackson, R. and Howe, N. [2003], *The 2003 Aging Vulnerability Index - An Assessment of the Capacity of Twelve Developed Countries to Meet the Aging Challenge*, The Center for Strategic and International Studies and Watson Wyatt Worldwide, Washington D.C. (USA).

Jagannathan, R. and Wang, Z. [1996], 'The conditional CAPM and the cross-section of expected returns', *Journal of Finance* **51**(1), 3–53.

Jermann, U. [1998], 'Asset pricing in production economies', *Journal of Monetary Economics* **41**(2), 257–275.

Jevons, W. S. [1871], *The Theory of Political Economy*, 3rd (1888) edn, Macmillan, London (UK).

Jäkel, M. F. [2002], New developments for the explanation of the equity premium puzzle. Unpublished Diploma Thesis, WHU - Otto Beisheim Graduate School of Management.

Jones, C. I. [1997], *Introduction to Economic Growth*, W.W. Norton & Company, New York NY (USA).

Jorgenson, D. W. [1963], 'Capital theory and investment behavior', *American Economic Review* **53**(2), 247–259.

Kaldor, N. [1963], Capital accumulation and economic growth, *in* F. A. Lutz and D. C. Hague, eds, 'Proceedings of Conference Held by the International Economics Association', Macmillan, London (UK).

Kamien, M. I. and Schwartz, N. L. [1991], *Dynamic Optimization: The calculus of variations and optimal control in economics and managements*, Vol. 31 of *Advanced Textbooks in Economics*, Elsevier, North Holland, Amsterdam (Netherlands).

Keynes, J. M. [1936], *The General Theory of Employment, Interest and Money*, Great Mind Series, Reprint by Prometheus Books [1997], New York NY (USA).

Kögel, T. [2005], 'Youth dependency and total factor productivity', *Journal of Development Economics* **76**(1), 147–173.

Kimball, M. S. [1990], 'Precautionary saving in the small and in the large', *Econometrica* **58**(1), 53–73.

Kinsella, K. and Phillips, D. R. [2005], 'Global aging: The challenge of success', *Population Bulletin - A publication of the Population Reference Bureau* **60**(1).

Kinsella, K. and Velkoff, V. A. [2001], *An Aging World: 2001*, International Population Reports, WP/02, U.S. Census Bureau, Government Printing Office, Washington D.C. (USA).

Kocherlakota, N. R. [1996], 'The equity premium: it's still a puzzle', *Journal of Economic Literature* **34**(1), 42–71.

Kogan, L. [2004], 'Asset prices and real investment', *Journal of Financial Economics* **73**(3), 411–431.

Koopmans, T. C. [1965], On the concept of optimal economic growth, *in* 'The Econometric Approach to Development Planning', North Holland, Amsterdam (Netherlands), pp. 225–300.

Kotlikoff, L. J. and Burns, S. [2004], *The Coming Generational Storm – What you need to know about America's economic future*, The MIT Press, Cambridge MA (USA).

Kotlikoff, L. J. and Summers, L. H. [1981], 'The role of intergenerational transfers in aggregate capital accumulation', *Journal of Political Economy* **89**(4), 706–732.

Krueger, D. and Kubler, F. [2001], Intergenerational risk sharing via Social Security when financial markets are incomplete. Working Paper, Stanford University.

Kydland, F. E. and Prescott, E. C. [1982], 'Time to build and aggregate fluctuations', *Econometrica* **55**(6), 1345–1370.

Lamont, O. A. [2000], 'Investment plans and stock returns', *Journal of Finance* **55**(6), 2719 – 2745.

Lee, R. D. [1994*a*], The formal demography of population aging, transfers, and the economic life cycle, *in* L. G. Martin and S. H. Preston, eds, 'Demography of Aging', National Academy Press.

Lee, R. D. [1994*b*], 'Population Age Structure, Intergenerational Transfer and Wealth - A New Approach, with Applications to the United States', *Journal of Human Resources* **29**(4), 1027–1063.

Lee, R. D. and Skinner, J. [1999], 'Will aging baby boomers bust the federal budget?', *Journal of Economic Perspectives* **13**(1), 117–140.

Leland, H. E. [1968], 'Saving and uncertainty: The precautionary demand for saving', *Quarterly Journal of Economics* **82**(3), 465–473.

Lengwiler, Y. [2004], *Microfoundations of Financial Economics - An Introduction to General Equilibrium Asset Pricing*, Princeton University Press, Princeton NJ (USA).

Lesthaeghe, R. and Neels, K. [2002], 'From the First to the Second Demographic Transition: An Interpretation of the Spatial Continuity of Demographic Innovation in France, Belgium and Switzerland', *European Journal of Population* **18**(4), 325–360.

Lesthaeghe, R. and Surkyn, J. [2004], When history moves on: The foundations and diffusion of a second demographic transition. Working Paper, Interface Demography, Vrije University.

Lesthaeghe, R. and van de Kaa, D. J. [1986], Twee demografische transities?, *in* R. Lesthaeghe and D. J. van de Kaa, eds, 'Bevolking: Groei en Krimp', Van Loghum Slaterus, Deventer (Netherlands), pp. 9–24.

Lettau, M. [2003], 'Inspecting the mechanism: The determination of asset prices in the RBC model', *Economic Journal* **113**(489), 550–575.

Li, Q., Vassalou, M. and Xing, Y. [2005], 'Sector investments growth rates and the cross-section of equity returns', *Journal of Business* . Forthcoming.

Liebman, J. B. [2001], Redistribution in the current u.s. Social Security system, *in* M. Feldstein and J. B. Liebman, eds, 'Distributional Aspects of Social Security and Social Security Reform', NBER Conference Report, University of Chicago Press, Chicago IL (USA), pp. 11–48.

Lim, K.-M. and Weil, D. N. [2003], 'The baby boom and the stock market boom', *Scandinavian Journal of Economics* **105**(3), 359–378.

Lindbeck, A. and Persson, M. [2003], 'The gains from pension reform', *Journal of Economic Literature* **41**(1), 74–112.

Lintner, J. [1965], 'The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets', *Review of Economics and Statistics* **47**(1), 13–37.

Lipsey, R. G. and Lancaster, K. [1956], 'The general theory of second best', *Review of Economic Studies* **24**(1), 11–32.

Lucas, R. E. [1967], 'Adjustment costs and the theory of supply', *Journal of Political Economy* **75**(4), 321–334.

Lucas, R. E. [1978], 'Asset prices in an exchange economy', *Econometrica* **46**(6), 1429–1445.

Lucas, R. E. [1988], 'On the mechanics of economic development', *Journal of Monetary Economics* **22**(1), 3–42.

Lucas, R. E. [2002], *Lectures on Economic Growth*, Harvard University Press, Cambridge MA (USA).

Lucas, R. E. and Prescott, E. C. [1971], 'Investment under uncertainty', *Econometrica* **39**(5), 659–681.

Mackenroth, G. [1952], Die Reform der Sozialpolitik durch einen deutschen Sozialplan, *in* E. Boettcher, ed., 'Schriften des Vereins für Sozialpolitik - Neue Folge', Vol. 4, Duncker und Humblot, Berlin (Germany), pp. 39–76.

Mackenroth, G. [1957], Die Reform der Sozialpolitik durch einen deutschen Sozialplan, *in* E. Boettcher, ed., 'Sozialpolitik und Sozialreform - Ein einführendes Lehr- und Handbuch der Sozialpolitik', Mohr, Tübingen (Germany), pp. 43–74.

Mackenzie, G. A., Gerson, P. R., Cuevas, A. and Heller, P. S. [2001], Pension reform and the fiscal policy stance. IMF Working Paper No. 01/214, International Monetary Fund.

Malliaris, A. G. and Brock, W. A. [1982], *Stochastic methods in economics and finance*, Vol. 17 of *Advanced Textbooks in Economics*, Elsevier, North Holland, Amsterdam (Netherlands).

Malthus, T. R. [1798], *An Essay on the Principle of Population, as it affects the Future Improvement of Society with Remarks on the Speculations of Mr. Goldwin, M. Condorcet, and other Writers*, Great Mind Series, Reprint by Prometheus Books [1998], New York NY (USA).

Malthus, T. R. [1815], 'An inquiry into the nature and progress of rent and the principles by which it is regulated', Reproduced from the 1815 edn in "The Pamphlets of Thomas Robert Malthus"[1970], London (UK).

Mankiw, N. G. and Weil, D. N. [1989], 'The baby boom, the baby bust, and the housing market', *Regional Science and Urban Economics* **19**(2), 235–258.

Mankiw, N. G. and Zeldes, S. P. [1991], 'The consumption of stockholders and non-stockholders', *Journal of Financial Economics* **29**(1), 97–112.

Markowitz, H. M. [1952], 'Portfolio selection', *Journal of Finance* **7**(1), 77–91.

Markowitz, H. M. [1991], *Portfolio Selection - Efficient Diversification of Investments*, 2nd (1991) edn, Basil Blackwell, Cambridege MA (USA).

Markowitz, H. M. [1999], 'The early history of portfolio theory: 1600-1960', *Financial Analysts Journal* **55**(4), 5–16.

Marshall, A. [1890], *Principles of Economics*, Great Mind Series, Abridged Reprint by Prometheus Books [1997], New York NY (USA).

Matsen, E. and Thogersen, O. [2004], 'Designing Social Security - a portfolio choice approach', *European Economic Review* **48**(4), 883–904.

Mayers, D. [1973], 'Nonmarketable assets and the determination of capital asset prices in the absence of a riskless asset', *Journal of Business* **46**(2), 258–267.

McMorrow, K. and Roeger, W. [1999], The Economic consequences of ageing populations - A comparison of the EU, US and Japan. Economic Paper No. 138 of the Directorate-General for Economic and Financial Affairs of the European Commission,.

Meadows, D. L., Meadows, D. H., Naill, R. F., Randers, J., Behrens, W. W. and Zahn, E. K. [1974], *Dynamics of growth in a finite world*, Productivity Press, Cambridge MA (USA).

Meadows, D. L., Meadows, D. H., Randers, J. and Behrens, W. W. [1972], *The Limits to Growth - A Report for the Club of Rome's Project*, Universe Books, New York NY (USA).

Mehra, R. [2003], 'The equity premium: why is it a puzzle', *Financial Analysts Journal* **59**(1), 54–69.

Mehra, R. and Prescott, E. C. [1985], 'The equity premium - a puzzle', *Journal of Monetary Economics* **15**(2), 145–161.

Merton, R. C. [1969*a*], 'An golden-rule for the welfare-maximization in an economy with a varying population growth rate', *Western Economic Journal* **4**, 307–318.

Merton, R. C. [1969*b*], 'Lifetime portfolio selection under uncertainty: The continuous-time case', *Review of Economics and Statistics* **51**(3), 247–257.

Merton, R. C. [1971], 'Optimum consumption and portfolio rules in a continuous-time model', *Journal of Economic Theory* **3**(4), 373–413.

Merton, R. C. [1973], 'An intertemporal capital asset pricing model', *Econometrica* **41**(5), 867–887.

Merton, R. C. [1975], 'An asymptotic theory of growth under uncertainty', *Review of Economic Studies* **42**(3), 375–393.

Merton, R. C. [1983], On the role Social Security as means for efficient risk-bearing in an economy where human capital is not tradable, *in* Z. Bodie and J. B. Shoven, eds, 'Financial Aspects of the United States Pension System', NBER Books, University of Chicago Press, Chicago IL (USA), pp. 325–358.

Merton, R. C. [1990], *Continuous-time Finance*, Revised edn. (Reprint 2003), Blackwell, Malden MA (USA).

Mirman, L. J. [1973], 'The steady state behaviour of a class of one-sector growth models with uncertain technology', *Journal of Economic Theory* **6**(1), 219–242.

Modigliani, F. [1986], 'Life cycle, individual thrift, and the wealth of nations', *American Economic Review* **76**(3), 297–313.

Modigliani, F. [1988], Measuring the contribution of intergenerational transfers to total wealth: Conceptual issues and empirical findings, *in* D. Kessler and A. Masson, eds, 'Modelling the Accumulation and Distribution of Wealth', Oxford University Press, Oxford (UK), pp. 21–52.

Modigliani, F. and Ando, A. [1963], 'The life cycle hypothesis of saving: Aggregated implications and tests', *American Economic Review* **53**(1), 55–64.

Modigliani, F. and Brumberg, R. [1954], Utility analysis and the consumption function: An interpretation of cross-section data, *in* K. Kurihara, ed., 'Post-Keynesian Economics', Rutgers University Press, New Brunswick NJ (USA).

Modigliani, F. and Miller, M. H. [1958], 'The cost of capital, corporate finance, and the theory of investment', *American Economic Review* **48**(3), 261–297.

Modigliani, F. and Muralidhar, A. [2004], *Rethinking Pension Reform*, Cambridge University Press, Cambridge (UK).

Mossin, J. [1966], 'Equilibrium in a capital asset market', *Econometrica* **34**(4), 768–783.

Mussa, M. [1977], 'External and internal adjustment costs and the theory of aggregate and firm investment', *Economica* **44**(174), 163–178.

Muth, J. A. [1961], 'Rational expectations and the theory of price movements', *Econometrica* **29**(6), 315–335.

Myles, G. D. [1995], *Public Economics*, Cambridge University Press, Cambridge (UK).

Nataraj, S. and Shoven, J. B. [2003], 'Comparing the risks of Social Security with and without individual accounts', *American Economic Review* **93**(2), 348–353.

Notestein, F. W. [1945], Population – the long view, *in* T. W. Schultz, ed., 'Food for the World', University of Chicago Press, Chicago IL (USA), pp. 36–57.

OECD [1998], *Maintaining Prosperity in an Ageing Society*, Organization for Economic Co-operation and Development, Paris (France).

OECD [2000], *Reforms for an Ageing Society*, Organization for Economic Co-operation and Development, Paris (France).

OECD [2004], 'Global pension statistics project: Measuring the size of private pensions with an international perspective', *Financial Market Trends* **2004**(87), 229–239. Organization for Economic Co-operation and Development.

Pareto, V. [1906], *Manual of Political Economy*, reprinted (1969) edn, Augustus M. Kelley, New York NY (USA).

Phelps, E. S. [1961], 'Golden rule of accumulation', *American Economic Review* **51**(4), 638–643.

Phelps, E. S. [1965], 'Second essay on the golden rule of accumulation', *American Economic Review* **55**(4), 793–814.

Porterba, J. M. [1998], Population age structure and asset returns: an empirical investigation. NBER Working Paper No. 6774, National Bureau of Economic Research.

Porterba, J. M. [2001], 'Demographic structure and asset returns', *Review of Economics and Statistics* **83**(4), 565–584.

Porterba, J. M. [2004], The impact of population aging on financial markets. NBER Working Paper No. 10851, National Bureau of Economic Research.

Pratt, J. W. [1964], 'Risk aversion in the small and in the large', *Econometrica* **32**(1), 122–136.

Pötzsch, O. and Sommer, B. [2003], *Bevölkerung Deutschlands bis 2050 - Ergebnisse der 10. koordinierten Bevölkerungsvorausberechnung*, Statistisches Bundesamt, Wiesbaden (Germany).

Raffelhüschen, B. [1999], 'Generational Accounting in Europe', *American Economic Review* **89**(2), 167–170.

Ramsey, F. P. [1928], 'A mathematical theory of saving', *Economic Journal* **38**, 543–559.

Restoy, F. and Rockinger, M. G. [1994], 'On stock market returns and returns on investment', *Journal of Finance* **49**(2), 543–557.

Ricardo, D. [1817], *Principles of Political Economy and Taxation*, Great Mind Series, Reprint by Prometheus Books [1996], New York NY (USA).

Robinson, J. [1938], 'The classification of inventions', *Review of Economic Studies* **5**(2), 139–142.

Rosen, H. S. [2002], *Public Finance*, 6th edn, McGraw-Hill, New York NY (USA).

Rosenzweig, M. R. [1990], 'Population growth and human capital investments: Theory and evidence', *Journal of Political Economy* **98**(5), 38–70.

Ross, S. A. [1976], 'The arbitrage theory of capital asset pricing', *Journal of Economic Theory* **13**(3), 341–360.

Ross, S. A. [1978], 'A simple approach to the valuation of risky streams', *Journal of Business* **51**(3), 453–475.

Rouwenhorst, G. K. [1995], Asset pricing implications of equilibrium business cycle models, *in* T. F. Cooley, ed., 'Frontiers of Business Cycle Research', Princeton University Press, Princeton NJ (USA), p. 294–330.

Roy, A. D. [1952], 'Safety first and the holding of assets', *Econometrica* **20**(3), 431–449.

Rubinstein, M. [1974], 'An aggregation theorem for securities markets', *Journal of Financial Economics* **1**(3), 225–244.

Rubinstein, M. [1976], 'The valuation of uncertain income streams and the pricing of options', *Bell Journal of Economics* **7**, 407–425.

Samuelson, P. A. [1947], *Foundations of Economic Analysis*, Harvard Economic Studies, enlarged (1983) edn, Harvard University Press, Cambridge MA (USA).

Samuelson, P. A. [1958], 'An exact consumption-loan model with or without social contrivance of money', *Journal of Political Economy* **66**(6), 467–482.

Samuelson, P. A. [1969], 'Lifetime portfolio selection by dynamic stochastic programming', *Review of Economics and Statistics* **51**(3), 239–246.

Samuelson, P. A. [1975*a*], 'The optimum growth rate for population', *International Economic Review* **16**(3), 531–538.

Samuelson, P. A. [1975*b*], 'Optimum Social Security in a life-cycle growth model', *International Economic Review* **16**(3), 539–5344.

Sandmo, A. [1970], 'The effect of uncertainty on saving decisions', *Review of Economic Studies* **37**(3), 353–360.

Schacht, U. [2001], *Finanztheoretische Analyse der Alterssicherung*, Europäische Hochschulschriften, Reihe 5, Volks- und Betriebswirtschaft, No. 2793, Peter Lang, Frankfurt/Main (Germany).

Scheinkman, J. A. [1989], Market incompleteness and equilibrium valuation of assets, *in* S. Bhattacharya and G. M. Constantinides, eds, 'Theory of Valutation: Frontiers of Modern Financial Theory', Vol. 1 of *Rowman and Littlefiels Studies in Financial Economics*, Rowman and Littlefiels, Totowa NJ (USA), chapter 2, pp. 45–51.

Schieber, S. J. and Shoven, J. B. [1994], The consequences of population aging on private pension fund saving and asset markets. NBER Working Paper No. 4665, National Bureau of Economic Research.

Schultz, T. W. [1961], 'Investment in human capital', *American Economic Review* **51(1)**, 1–17.

Schultz, T. W. [1963], *The Economic Value of Education*, Columbia University Press, New York NY (USA).

Schultz, T. W. [1974], *Marriage, Human Capital and Fertility*, University of Chicago Press, Chicago IL (USA).

Sethi, S. P. and Thompson, G. L. [2000], *Optimal Control Theory: Applications to management science and economics*, 2nd edn, Kluwer Academic, Norwell MA (USA).

Sharpe, W. F. [1964], 'Capital asset prices: A theory of market equilibrium under conditions of risk', *Journal of Finance* **19**(3), 425–442.

Shell, K. [1971], 'Notes on the economics of infinity', *Journal of Political Economy* **79**(5), 1002–1011.

Shiller, R. J. [1982], 'Consumption, asset markets, and macroeconomic fluctuations', *Carnegie-Rochester Conference Series on Public Policy* **17**, 203–238.

Shiller, R. J. [1993], *Macro Markets: Creating Institutions for Managing Society's Largest Economic Risks*, Clarendon Lectures in Economics, Oxford University Press, Oxford (UK).

Shiller, R. J. [1998], Social Security and institutions for intergenerational, intragenerational and international risk sharing. Working Paper presented at the Carnegie–Rochester Public Policy Conference.

Shiller, R. J. [2003*a*], *The New Financial Order: Risk in the 21st Century*, Princeton University Press, Princeton NJ (USA).

Shiller, R. J. [2003*b*], 'Social Security and individual accounts as elements of overall risk-sharing', *American Economic Review* **93**(2), 343–347.

Siegel, J. J. [1998], *Stocks for the Long Run*, 2nd edn, McGraw-Hill, New York NY (USA).

Simon, J. and Zinsmeister, K. [n.d.], Eerdmans, Title = How Population Growth Affects Human Progress, Year = 1995, Keywords = Population and Economy, Grand Rapids MI (USA), chapter 3, pp. 61–79.

Smith, A. [1776], *Wealth of Nations*, Great Mind Series, Reprint by Prometheus Books [1991], New York NY (USA).

Solow, R. M. [1956], 'A contribution to the theory of economic growth', *Quarterly Journal of Economics* **70**(1), 65–94.

Solow, R. M. [1957], 'Technical change and the aggregate production function', *Review of Economics and Statistics* **39**(3), 312–320.

Solow, R. M. [1969], Investment and technical change, *in* K. J. Arrow, S. Karlin and P. Suppes, eds, 'Mathematical Methods in the Social Science', Stanford University Press, Palo Alto CA (USA), pp. 89–104.

Stambaugh, R. F. [1983], 'Arbitrage pricing with information', *Journal of Financial Economics* **12**(3), 357–370.

Storesletten, K., Telmer, C. I. and Yaron, A. [2001], Asset pricing with idiosyncratic risk and overlapping generations. Seminar Paper No. 703, Institute for International Economic Studies, Stockholm University.

Storesletten, K., Telmer, C. I. and Yaron, A. [2004], 'Consumption and risk sharing over the life cycle', *Journal of Monetary Economics* **51**(3), 609–633.

Sundaresan, S. M. [1984], 'Consumption and equilibrium interest rates in stochastic production economies', *Journal of Finance* **39**(1), 77–92.

Swan, T. W. [1956], 'Economic growth and capital accumulation', *Economic Record* **32**, 334–361.

Taylor, A. [1996], International capital mobility in history: The saving-investment relationship. NBER Working Paper No. 5743, National Bureau of Economic Research.

Thompson, W. S. [1929], 'Population', *American Journal of Sociology* **35**, 959–975.

Tobin, J. [1958], 'Liquidity preference as behavior towards risk', *Review of Economic Studies* **25**(1), 65–86.

Tobin, J. [1969], 'A general equilibrium approach to monetary theory', *Journal of Money, Credit and Banking* **1**(1), 15–29.

Turner, D., Giorno, C., De Serres, A., Ann, V. and Richardson, P. [1998], The macroeconomic implications of ageing in a global context. Working Paper No. ECO/WKP(98)6, Organization for Economic Co-operation and Development.

United Nations [1999], *The World at Six Billion*, Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York (United Nations).

United Nations [2000], *Replacement Migration: Is it A Solution to Declining and Ageing Populations?*, Population Division of the Department of

Economic and Social Affairs of the United Nations Secretariat, New York (United Nations).

United Nations [2001], *World Population Aging: 1950-2050*, Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York (United Nations).

United Nations [2004*a*], *World Population to 2300*, Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York (United Nations).

United Nations [2004*b*], *World Urbanization Prospects, The 2003 Revisions*, Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York (United Nations).

United Nations [2005*a*], *World Population Prospects: The 2004 Revision, Highlights*, Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York (United Nations). Forthcoming.

United Nations [2005*b*], *World Population Prospects: The 2004 Revision, Vol. I, Comprehensive Tables*, Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York (United Nations). Forthcoming.

United Nations [2005*c*], *World Population Prospects: The 2004 Revision, Vol. II, Sex and Age Distribution of the World Population*, Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York (United Nations). Forthcoming.

United Nations [2005*d*], *World Population Prospects: The 2004 Revision, Vol. III, Analytical Report*, Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, New York (United Nations). Forthcoming.

U.S. Census Bureau, ed. [2004], *Global Population Profile: 2002*, International Population Reports, P95/01-1, U.S. Census Bureau, Government Printing Office, Washington D.C. (USA).

Uzawa, H. [1961], 'Neutral inventions and the stability of growth equilibrium', *Review of Economic Studies* **28**(2), 117–124.

Uzawa, H. [1965], 'Optimal technical chance in an aggregative model of economic growth', *International Economic Review* **6**(1), 18–31.

Valdés, B. [1999], *Economic Growth - Theory, Empirics and Policy*, Edward Elgar, Cheltenham (UK).

Vaupel, J. W. [1998], 'Demographic analysis of aging and longevity', *American Economic Review* **88**(2), 242–247.

Viceira, L. M. [2001], 'Optimal portfolio choice for long-horizon investors with nontradeabel labor income', *Journal of Finance* **56**(2), 433–470.

von Hayek, F. A. [1941], *The Pure Theory of Capital*, University of Chicago Press, Chicago IL (USA). Reprint 1971.

von Neumann, J. and Morgenstern, O. [1944], *Theory of Games and Economic Behavior*, 60th anniversary commemorative (2004) edn, Princeton University Press, Princeton NJ (USA).

von Weizäcker, J. [1999], Mixing pay-as-you-go and funded pension systems. Working Paper, University or Munich.

Walras, A. [1874], *Elements of Pure Economics: Or the theory of social wealth*, Richard Irwin, Homewood IL (USA). Translation (1956) of revised edition (1926).

Weil, P. [1989*a*], 'Confidence and the real value of money in an overlapping generations economy', *Quarterly Journal of Economics* **102**(1), 1–22.

Weil, P. [1989*b*], 'The equity premium puzzle and the risk-free rate puzzle', *Journal of Monetary Economics* **24**(3), 401–421.

Weil, P. [1989*c*], 'Overlapping families of infinitely-lived agents', *Journal of Public Economics* **38**(1), 183–198.

Weir, D. R. [1987], Malthus's theory of population, *in* J. Eatwell, M. Milgate and P. Newman, eds, 'The New Palgrave: Economic Development', Macmillan, London (UK), pp. 226–231.

Wheelan, C. [2002], *Naked Economics: Undressing the Dismal Science*, W.W. Norton & Company, New York NY (USA).

Williams, J. T. [1978], 'Risk, human capital, and the investor's portfolio', *Journal of Business* **51**(1), 65–89.

Willis, R. J. [1987], 'What have we learned from the economics of the family?', *American Economic Review* **77**(2), 68–81.

Wilson, R. [1968], 'The theory of syndicates', *Econometrica* **36**(1), 119–132.

Winkelmann, R. and Zimmermann, K. F. [1992], Ageing, migration and labour mobility. C.E.P.R. Discussion Paper No. 706, Centre for Economic Policy Research.

World Bank [1994], *Averting the Old Age Crisis: Policies to Protect the Old and Promote Growth*, (A World Bank Policy Research Report), Oxford University Press, Oxford (UK). The International Bank for Reconstruction and Development.

Yoo, P. S. [1997], Population growth and asset prices. Working Paper No. 1997-016A, Federal Reserve Bank of St. Louis.

Zhang, L. [2005], 'The value premium', *Journal of Finance* **60**(1), 67–104.