PENSIONS REFORMS, REDISTRIBUTION AND WELFARE

by

JEEVENDRANATH (VIMAL) THAKOOR

A thesis submitted to
The University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

Department of Economics
School of Social Sciences
The University of Birmingham
May 2009

Dedicated to my parents,
Arvind & Tara Thakoor
for everything...

Table of Contents

Abstract

This thesis deals with the optimal design of pensions systems in the face of demographic changes. Though the chapters differ in terms of the key questions addressed, the unifying theme remains which pensions system yields the highest welfare under differing economic conditions. We use a standard overlapping generations model with heterogeneous agents to address the various questions. The role of the pensions system varies between consumption smoothing and redistribution, or a combination of both. The provision of pensions, whether universal or targeted, has a significant impact on capital formation and by extension on a host of economic aggregates and welfare. Capital is always higher under a fully-funded scheme. Under certain conditions, it is optimal to have no pay-as-you-go pensions in place and a fully-funded scheme is thus optimal. With a redistributive pensions system, the welfare gain of the poor exceeds the fall in the welfare of the rich thereby resulting in an increase in aggregate welfare. This thesis thus brings together the issues involved in pensions design in a theoretical framework and aims to provide an insight into the various channels at work.

Keywords: *Pensions; Ageing Population, Pensions Reforms; Pay-As-You-Go Pensions; Fully Funded Pensions;Redistribution; Income Support; Overlapping Generations Model (OLG); General Equilibrium; Heterogeneous Agents; Dynamic Efficiency; Dynamic Inefficiency; Welfare.*

Acknowledgements

During my Ph.D. I have benefited from the advice, support and encouragement of many people.

My supervisor, Jayasri Dutta, has been a major influence throughout my time at Birmingham. She was always available, often beyond the call of duty. I also benefitted from significant discussions with Herakles Polemarchakis on various aspects of my work. Their influence is woven throughout this thesis and I owe them a huge debt of gratitude.

Colin Rowat provided constructive comments and advice on my work during the early stages of the thesis. Indra Ray provided invaluable support and guidance.

Engin Kara deserves my thanks for his interest in my research and help during his time at Birmingham.

This final version of the thesis has benefitted immensely from the insightful comments and suggestions from the members of my thesis committee, Professors Sayantan Ghosal and Peter Sinclair.

The various seminar participants in Birmingham, Salamanca (Spain), Utrecht (Netherlands), the Institute for Fiscal Studies (London), and the Bank of Mauritius provided valuable ideas.

I gratefully acknowledge the financial support from the Department of Economics. A special thanks to the support staff in the Department for all their help, often at short notice.

My friends provided a welcome distraction away from work. I have fond memories of time spent with Andreas Olfert, Ingrid Fiedler, Kavin Hemraj, Nicolas Martin, Nicholas Vassilakos, Simone Gobien, Susanne Quadros, and Wei Chen.

Priscilla Muthoora has been a great friend and companion through this journey, providing unflinching support and motivation. She encouraged me to apply for the Chevening Scholarship, which helped fund the first year of my Ph.D., and patiently discussed my nascent ideas, many of which are now in the thesis.

Last but not least, I would like to express my gratitude to my family, beginning with my grandparents, for their advice and inspiration. Mum and dad have always supported me selflessly in all my endeavours, and made everything possible. My brother Kamal was a constant source of support and kept me going. A special word for Mr. and Mrs. Hemraj for opening their home to me during my time in England.

1 INTRODUCTION

The post-war baby boom coupled with an increasing life expectancy and declines in fertility rates have resulted in the ageing of the population of many countries. While it was initially a phenomenon restricted to industrial nations, developing countries have also started to experience this transition, and at a relatively faster pace. An ageing population poses a number of policy conundrums for policymakers. These span the provision of pensions and health care coupled with implications for the macroeconomy and financial stability. One of the prominent challenges of an ageing population is the provision of pensions for the elderly so as to ensure they have a sufficient amount of income during their retirement. Fears about the long term fiscal-financial sustainability of pensions schemes operated under the Pay-As-You-Go (PAYG)[1] have been raised as the dependency ratio increases. Without reforms, some of the PAYG schemes are expected to go bust. While it makes economic sense to reform unsustainable systems, attempts for reform have often encountered severe (political) resistance. Though the initial policy

[1] In a PAYG pensions system, there is intergenerational transfers such that the current cohort of workers pay for the pensions of the current cohort of retirees. In a FF system, each generation save for its own retirement.

proposals suggested a shift from the PAYG to a Fully-Funded (FF) system, the transition has be hindered by several factors. This includes the high transition cost of shifting from the PAYG to a FF system. Moreover, most developing countries do not have the capabilities to manage such a transition on their own. There is also the no less important issue of coverage whereby people living in the rural areas or working in the informal sector would not be covered by the FF system. As such, they risk being in poverty during their retirement. One of the main perceived benefits of a FF system is that the assets are privately managed and yield a higher rate of return than the PAYG. However, this view has come under challenge, especially in light of the recent developments in the financial markets which has resulted in pensions funds losing a significant proportion of their portfolio and thereby jeopardising the income of a cohort of retirees. Moreover, there is an ongoing debate about whether the rate of return of the FF is no more higher than the PAYG once the risk elements have been taken into account.

In a bid to prepare their economies to cope better with the ageing of their population, most countries have initiated, or thought of initiating, the reforms of their pensions system. These reforms have been parametric or fundamental. The most well known example of fundamental reform is that of Chile, which was the first country to move from a PAYG to a FF system in 1981. Since then, there has been a wave of pensions reform which has swept through Latin America. However, fundamental reform often encounters significant resistance. To circumvent this problem, most countries have emphasised parametric reforms whereby the PAYG pensions is maintained but to deal with ageing, the retirement age is extended or the generosity of the system reduced. There is now a general consensus that an optimal pensions system will have a combination of both the PAYG and the FF schemes. This will ensure a better diversification of risks that could potentially arise if only one system was adhered to.

Pensions were initially introduced in the US and Europe to ensure that the elderly do not live in poverty. This was based on paternalistic and equity considerations since some individuals would not be able to save enough for their retirement. Though pensions was initially provided universally, this cannot be sustained with an ageing population. As such, the question of equity versus efficiency arises. There is an emerging view that a "means-tested" policy, whereby only the poor benefit most from pensions, is more resilient to ageing than one where pensions is universally provided. Moreover, though the paternalistic motive for the provision of pensions should still be there, a pensions policy based on targeting will be more efficient. An ageing population thus has major ramifications for pensions policy and the overall economy.

In this thesis "Pensions Reforms, Redistribution and Welfare" we focus on the optimal design of pensions systems. Though the chapters differ in terms of the key question addressed, the unifying theme throughout remains which pensions system yields the highest welfare under differing economic conditions. The role of the pensions system varies between consumption smoothing and redistribution or a combination of both. The differences in the design of the pensions system lead to differences in a range of economic aggregates and by extension welfare. One of the key channels through which this operates is the difference in capital formation. In a FF system capital is always higher than a PAYG scheme. This has implications for other aggregates such as output, wages, consumption and the interest rate which all play an important role in determining the welfare of the different agents. This thesis brings together the issues involved in pensions design in a theoretical framework and aims to provide an insight into the various channels at work.

This thesis consists of 4 core chapters. We use a standard Samuelson-Diamond overlapping generations model initiated by Samuelson (1958) and extended by Diamond (1965) to address the questions posed. At any point in time, two generations, the young and old live simultaneously. Agents maximise utility by maximising consumption and for ease of manipulation, the utility function is assumed to be loglinear. Abstracting from any adverse impact of social security we assume that the young provide one unit of labour inelastically. The old live in retirement. We allow the agents to differ through a combination of either myopia and/or productivity. The high (low) productivity agents are termed as rich (poor). In line with Becker (1990), we assume that the poor can potentially discount the future at a higher rate than the rich. To allow for the potential adverse impact of pensions on savings, we allow for endogenous capital formation. To complete the general equilibrium set up, the economy consists of profit maximising firms and a welfare maximising planner. Firms produce a homogeneous good using a Cobb-Douglas production function. The analysis is carried out in steady state and as such we do not consider the transition cost in switching from a PAYG to a FF scheme. Following Galor (1992), the steady state can be considered as the representative framework within which infinitely many generations evolve. The key question deals with how the agents fund for their consumption in retirement. Some of the possibilities we consider are: (i) pensions is FF and all agents are responsible for the provision of their own pensions ($Ch.$ 2-4); (ii) pensions is in the form of intergenerational transfers (PAYG) ($Ch.$ 2-5); (iii) there is a multipillar pensions system in place where part is FF and part operates under PAYG ($Ch.$ 3-5). The part that operates under PAYG also has an element of redistribution from the rich to the poor ($Ch.$ 5). In chapters 3-5, we estimate some parameter values for a sample of countries, developed and developing, and try to find out what it the optimal level of tax that would be imposed in those economies. Since OLG models are inherently hard to calibrate to real world data, we also use sensitivity analysis to show how the tax rate behaves as the parameter values change. The results remain robust for a whole set of plausible parameters.

Chapter 2 provides a comparison between a FF and a PAYG scheme in a general equilibrium framework. This is an extension of a paper by Feldstein (1985) who finds a positive optimal level of social security aimed at smoothing consumption in a dynamically efficient economy characterised by myopia. This is counter to the Aaron condition (1966) which builds on work by Samuelson (1958) and Diamond (1965). Samuelson (1958) proved the existence of an optimal "biological" interest rate equal to the population growth rate and Diamond (1965) showed that it is possible for there to be overaccumulation of capital in OLG models thereby leading to dynamic inefficiency. The Aaron condition suggests that if the interest rate is greater than the growth of the population and wage rate, then a PAYG leads to a further decline in capital and is not optimal. Feldstein's work relies on a fixed rate of interest of 11.4 percent. However, this is counter to Lerner's (1959) criticism of Samuelson's constant interest rate. Lerner suggests that the interest rate will in fact vary from one period to another based on the extent of pensions – a higher (lower) pensions at retirement will encourage (discourage) people to consume more (less) during their working years and thereby increase (decrease) the interest rates. However, most of the criticisms of the PAYG initially started from the Chicago School with Friedman (1962) one of the most ardent critiques suggesting social security programmes are inappropriate because they infringe individual liberty.

We extend Feldstein's (1985) paper in two ways: Firstly we allow for endogenous capital formation to take into account the distortionary impact of a PAYG on savings and secondly, we compare the PAYG with a FF scheme to see which is the optimal pensions system. Consistent with Feldstein (1985), the agents suffer from myopia and this hinders their ability to give a sufficient weight to the future and fully anticipate the amount of pensions they will receive. Whilst we derive analytical results we also have recourse to simulations to show the full general equilibrium effects of the pensions systems.

Our results on PAYG are consistent with Feldstein (1985) and we show that capital is always higher under the FF system than the PAYG. This is in line with previous empirical results by Feldstein (1974, 1996). When myopia affects the expected amount of pensions to be received, the crowding is only partial and there is a convergence between capital under the two schemes.

Feldstein (1985) showed the existence of a positive level of PAYG pensions. However, for all the sets of simulations we undertake, the results show that, in a dynamically efficient economy, a FF system always yields a higher level of welfare than the PAYG. This is consistent with the Aaron (1966) condition. Moreover, the greater the myopia and the less the weight the agents attach to the second period of their lifetime, the higher the tax rate. This is in line with consumption smoothing. However, if myopia leads agents to expect a smaller pensions than they actually receive, there is a fall in the tax rate and convergence in welfare between the PAYG and FF systems. Nonetheless, the FF system remains optimal for all positive PAYG taxes. We also find that beyond a certain level of myopia, it is optimal to have no PAYG pensions in place. In such cases, a FF scheme is the only option.

One of the reasons pensions was initially provided was to ensure that the elderly do not live in poverty during retirement. However, with population ageing, there have been calls to "means-test" the provision of PAYG pensions. As such, only the elderly will receive such a payment and the rich will have to fund for their own pensions. Pensions would thus provide some form of social insurance and help alleviate poverty and/or reduce inequality. This is the crux of **Chapter 3**. The agents are heterogeneous both in terms of their productivity and the weight they attach to the future. We consider what is the optimal way to redistribute from the rich to the poor, i.e., whether it should be in the form of income support or PAYG pensions. If redistribution is intragenerational, a FF pensions system is in place. However, if there is intergenerational redistribution a hybrid system is in place where part is FF and part operates as a PAYG.

The issue of redistribution is not new and remains controversial. A growing literature has emphasised the various channels through which inequality can affect politico-economic stability (Persson and Tabellini (1994); Alesina and Rodrik (1994); Alesina and Perrotti (1996)). Redistribution is often considered as one of the routes through which social justice and efficiency can be promoted by reducing inequality and supporting those at the lower end of the economy. Though intragenerational redistribution has been the main tool, the redistributive role of pensions is gaining increasing prominence (Krueger and Kubler (2006)). Our paper comes closest to Conde-Ruiz and Galasso (2005) who find that with sufficient inequality in earnings and elderly in the economy, there is an equilibrium that supports the existence of both intra and intergenerational redistribution.

Our theoretical analysis provides further evidence of the distortionary impact of redistribution on capital. With income support, the impact on capital is smaller than with pensions. This is consistent with our earlier result pertaining to capital being higher under the FF scheme than a PAYG. Our results on intragenerational transfers suggest that a richer and more equal economy, characterised by a high proportion of rich and productivity, will require a smaller tax rate than an economy characterised by high inequality. The poor prefer higher taxes since this means the redistribution is more generous whilst the rich prefer lower taxes. In the same line, we find that welfare is higher in a richer economy.

Our simulations results suggest that although redistribution is costly, it almost always leads to an increase in welfare except for a small range of parameter values where the feasible optimal tax rate is zero. Redistribution increases aggregate welfare because the gain in the utility of the poor is higher than the loss to the utility of the poor. We also find that there is a potential for dynamic inefficiency to arise in the economy. This becomes a possibility when there is a high proportion of rich or productivity is fairly high. Under those circumstances, we find that redistribution through pensions is optimal. Redistributing through pensions in a dynamically inefficient economy leads to a lower capital and can move the economy to a dynamically efficient position. On the other hand, if the economy is in a dynamically efficient position, then income support is the preferred redistribution instrument. An economy is more likely to be dynamically efficient if there is a high proportion of poor agents with low productivity. These results are consistent with the Aaron condition and remain robust to population ageing. However, there is a small range of parameters in a dynamically efficient economy where the interest rate is greater than but very close to the population growth rate. In such a situation, pensions is an optimal redistribution instrument even in a dynamically efficient economy. We suggest the PAYG pensions ensure the economy remains dynamically efficient.

Chapter 4 addresses the same question as in Chapter 3, i.e., how to redistribute from the rich to the poor. However, there are two fundamental differences: (i) the planner has the option of redistributing both intergenerationally and intragenerationally; and (ii) we allow for the planner to potentially have some degree of inequality aversion whereby he gives a higher weight to the welfare of the poor. Additionally, the agents have the same discount factor and differ only in terms of their productivity. We investigate what determines which proportion is allocated to the young in the form of income support and which proportion goes to the elderly in the form of pensions. If everything is allocated to the young, we have a FF system, otherwise a hybrid system emerges. This paper is thus analogous to Conde-Ruiz and Galasso (2005) who investigated both inter and intragenerational redistribution simultaneously.

The analytical results are in line with those in Chapter 3. We find that the adverse impact of redistribution on capital still persists. The higher the intergenerational redistribution, the greater the crowding out. Moreover, we also find that richer economies consisting of a high proportion of rich and productivity need less redistribution and therefore lower taxes. The higher the inequality aversion of the planner, the greater the extent of redistribution and hence the higher the tax rate.

To consider the general equilibrium impact of the two instruments simultaneously, we resort to simulations. We find that the tax rate and the timing of redistribution change in such a way so as to ensure that capital does not change significantly and the economy remains dynamically efficient. We note that an increase in the tax rate would lead to a move in favour of intragenerational redistribution. Conversely, in a rich economy characterised by a high proportion of rich and productivity, the tax rate would tend to be low and intergenerational transfers would be favoured. When the agents suffer from myopia, intragenerational transfers are favoured since this reverses some of the decline in capital that is induced by myopia. We also find that with population ageing, intergenerational transfer is optimal. Consistent with our previous results, we find that both inter and intragenerational redistribution are supported only within a range. Outside this range, only one of the two is favoured. If an economy is poor intragenerational redistribution will be favoured whilst in a rich economy intergenerational redistribution will be optimal.

Finally, in **Chapter 5** we consider a multi-pillar pensions system. The predominant view that prevailed through most of Latin America after the privatisation of the Chilean transition to a FF scheme has changed drastically since the publication of the World Bank's "Averting the Old Age Crisis"

(1994). According to that report, a pensions system would ideally have three pillars so as to diversify the risks of both the FF and PAYG as well as ensuring all individuals, especially the poor, are catered for in retirement. It has to be acknowledged that "Averting the Old Age Crisis" has had its fair share of criticism. For instance, Gillion et al. (2000) suggest there are more reform options that are possible, than just the ones suggested by the Bank. Even the Bank in its subsequent work has pointed out that issues of coverage and the management of the privatised pensions funds had to be given due attention (Gill et al. (2004); World Bank (2005)). They thus suggested the design of a five-pillar pensions system with greater flexibility to adjust to different economic environments.

In Chapter 5 we consider a three-pillar pensions system whereby we aim to formalise the intuition behind the World Bank model. We consider some of the redistributive issues and the channels through which the three-pillars affect the macroeconomic aggregates and welfare. The agents in this model differ both in terms of their productivity and discount factor. The main difference with the models in the previous chapters is that pensions is provided to all the agents and there are three pillars. Pillar 1 is entirely redistributive and can be considered as a Beveridgean system since the pensions the agents get is independent of their contribution. Pillar 1 thus promotes an element of intragenerational redistribution. Pillar 2 is Bismarckian in that the pensions the agents get is a function of their contribution rate (which is equal to their ability). We differ from the World Bank in that we assume the same rate of return on both Pillars 1 and 2. However, as we have discussed earlier, there is no reason to assume that a privately managed Pillar 2 will undoubtedly yield a higher return than a publicly managed Pillar 1. Pillar 3 is entirely voluntary in this set up and it represents the savings the agents undertake irrespective of Pillars 1 and 2. The weight the planner attaches to Pillar 1 determines the extent of redistribution that takes place through pensions. The poor will favour a higher weight on Pillar 1 whilst the rich will favour Pillar 2.

The analytical findings remain consistent in so far as the redistributive impact of Pillar 1 is concerned. Capital is lower the higher the weight attached to Pillar 1. The higher the weight attached to Pillar 1, the higher (lower) the welfare of the poor (rich). Welfare is higher in a richer economy resulting from a combination of either higher proportion of rich and/or productivity. We also find that population ageing leads to a marginally lower level of welfare.

The key simulations results suggest that, for some plausible range of parameters, whether the planner decides to attach a higher weight to Pillar 1 or to Pillar 2 makes a marginal difference to the optimal tax rate. Our other simulations results suggest an increase in the weight attached to Pillar 1 leads to a *marginally* higher tax rate (and lower capital). However, since the welfare of the poor increases by more than that of the rich, aggregate welfare increases. We also find that richer economies characterised by a high proportion of rich are able to afford more generous pensions and have a higher welfare. However, the impact of productivity is non-linear, though a high levels of productivity, a higher pensions can be paid out. We suggest that the increase in the tax rate as the economies get richer ensures the economy remains dynamically efficient. Though the impact of population ageing on the tax rate is infinitesimal, it leads to a lower welfare.

2 OPTIMAL SOCIAL SECURITY - A GENERAL EQUILIBRIUM APPROACH

2.1 Introduction

The optimal level of social security has attracted significant academic interest since Samuelson (1958). Much is based on the premise that if an economy is dynamically efficient such that the prevailing rate of interest is greater than the population growth rate, then a fully-funded (FF) pensions system is optimal. On the other hand, if the population growth rate exceeds the interest rate, then a pay-as-you-go (PAYG) pensions system can be welfare improving. This concept is formalised in the Aaron condition whereby Aaron (1966) suggests that if the growth rate of population and real earnings per head exceeds the market rate of interest, then a PAYG system yields a higher welfare than a FF system. However, in an influential paper, Feldstein (1985) shows the existence of an optimal level of PAYG social security in a dynamically efficient economy characterised by myopia. Pensions thus aim to achieve consumption smoothing.

In this paper, we aim to extend[2] Feldstein's paper in two ways: Firstly,

[2] Whilst we try to adhere to the parameters used by Feldstein as closely as possible, we

Feldstein does not consider capital formation and his simulations results are based on an exogenous interest rate of 11.4 percent, which is the marginal product of capital in the US for the period 1950-1980. However, this potentially has two problems in that the general equilibrium effects are not accounted for and taking a fixed interest rate misses out on the distortions induced in capital formation when a PAYG pensions is provided. We include capital formation and endogenise the interest rate to overcome these problems. Secondly, for comparative purposes, we include a FF system as well. This enables us to consider which of the FF or PAYG deliver higher welfare. Whilst we are able to derive some analytical results, consistent with Feldstein we also have recourse to simulations to show the full general equilibrium effects of the pensions systems.

Like Feldstein, we use a Samuelson-Diamond overlapping generations model where two agents, young and old live simultaneously. The agents differ according to their level of myopia. The myopia can take two forms: firstly, it reduces the weight the agents attach to the second period of their

divert in some instances to attribute the parameters their more conventional usage. Feldstein uses α as a myopia parameter. In this paper, α is the share of capital in production. Feldstein's α is now replaced with σ. A_t is used for the number of retirees at time t, here we use it for technology at time t.

life and thus fail to save enough. Secondly, the agents may not be able to fully anticipate the amount of pensions they are going to receive. The economy also consists of profit-maximising perfectly competitive firms and a welfare-maximising planner.

The analytical results in so far as the PAYG is concerned are consistent with Feldstein. The key findings can be summarised as:

1. Capital is always higher under the FF system than the PAYG;

2. Whilst PAYG pensions causes crowding out, if myopia also affects the expected level of pensions to be received, then the crowding out is not complete;

3. A reduction in myopia leads to an increase in capital whilst an increase in the tax rate or the rate of population growth leads to a fall in capital.

Feldstein showed the existence of a positive level of PAYG pensions. However, once we take into account the impact of pensions on capital formation, the simulations results show that in a dynamically efficient economy a FF system always yields a higher welfare than the PAYG. This is consistent with the Aaron condition. The other simulations results can be summed up as:

1. The findings on capital remain as before and capital remains higher under the PAYG;

2. The higher the myopia, the higher the tax rate. This is consistent with Feldstein and consumption smoothing;

3. If myopia leads agents to expect smaller pensions than they actually receive, there is a convergence in welfare between the PAYG and FF systems. However, in such a situation, as myopia increases, the tax rate falls. The FF system still remains optimal for all positive PAYG taxes;

4. Consistent with Feldstein, beyond a certain point it is optimal to have no pensions in place.

The rest of this paper is structured as follows: in Section 2 we describe the set up of the economy. Section 3 derives the competitive equilibrium whilst Section 4 considers the planner's problem in terms of finding the optimal level of social security and considers a set of simulations. Section 5 concludes.

2.2 The Economy

The economy consists of profit-maximising firms operating in competitive markets, utility-maximising agents and a welfare-maximising benevolent planner. Economic activity takes place over infinite discrete time $t \in \{0, 1..., \infty\}$. At time t, there are two generations in place, the young and the old. The young provide are economically active whilst the old live in retirement. The agents are similar in all other aspects except for the degree of their myopia. Population grows at a constant rate n such that at any point in time there are $(1+n)$ more workers than retirees. This can be expressed as:

$$L_t = (1+n)L_{t-1} \tag{1}$$

where L_t is the number of agents born at time t. Technology grows at a constant rate g and hence

$$A_t = (1+g)A_{t-1} \tag{2}$$

where A_t is the technology prevailing in the current period.

The economy also consists of profit maximising perfectly competitive firms and a welfare maximising social planner. The planner maximises the welfare of all agents born at time t. Two factors, an amount of capital (k) and labour (l) are available as inputs to production and a homogeneous good (y) is produced.

2.2.1 Households

At time t, two generations live simultaneously. In line with Samuelson (1958), the distribution of the population is considered to be stationary such that the proportions and types of individuals remain the same across generations. Each young agent is endowed with one unit of labour and, abstracting from the potentially negative impact of social security on labour supply, we assume that each young agent supplies one unit of labour inelastically. Agents differ only in terms of their myopia which hinders their ability to anticipate their retirement and potentially save enough for the second period of their lifetime. Myopia may also result in agents not being able to fully anticipate the amount of pensions they are going to receive. In line with Feldstein, we assume that the population consists of a component of "life-cyclers", that is, those who base their economic decision on the two periods of their lifetime, and the remainder are myopes.

Agents are utility-maximising and utility is derived out of consumption

$u(c_t, c_{t+1})$ in the two periods of lifetime. We abstract from bequests such that people are born without any initial asset endowment, other than their labour supply, and they consume all the income they generate within their lifetime such that they bequeath nothing to the next generation. Agents thus choose their levels of consumption and savings to maximise their utility.

In line with Feldstein, we assume that utility derived out of maximising consumption subject to the budget constraint can be expressed additively such that: $U = u(c_1) + v(c_2)$ where U represents total utility over the lifetime of the agent and $u(c_1)$ and $v(c_2)$ represent the utility derived out of consumption in the first period and second periods of life respectively. For ease of manipulation, the intertemporally additive lifetime utility function is taken to be log-linear and satisfies all the usual conditions in the form of strict concavity such that $u'(c) > 0$ and $u''(c) < 0$. The function also satisfies $\lim_{c \to 0} u'(c) = \infty$ such that subject to its disposable income, the household will always choose a positive level of consumption when maximising life-cycle utility. The agents' problem can be thus expressed as:

$$\underset{\{c_t^y, c_{t+1}^o, s_t\}}{Max} : U = \ln c_t^y + \lambda^h \ln c_{t+1}^o \tag{3}$$

subject to:

$$c_t^y = w_t^y(1-\theta) - s_t^y \tag{4}$$

$$c_{t+1}^o = R_{t+1}s_t^y + \sigma^h b_{t+1} \tag{5}$$

Eqn.(3) represents the agents maximisation problem pertaining to consumption in the two periods of his lifetime. λ^h represents the level of myopia such that an individual with $\lambda = 1$ is a life cycler who values all periods the same, whilst an individual with $\lambda = 0$ values consumption only in the first period . For the vast majority of agents, $\lambda \in (0,1)$ such that at least a positive weight, small or large, is given to consumption in the second period of lifetime.

Eqns. (4)-(5) represent the consumption of the individual during the two periods of lifetime. c_t^y is the level of consumption in the first period suggesting that any disposal income, after the payment of a proportional tax θ, is allocated *à la* Diamond (1965) between present consumption and savings (s). The second period consumption, c_{t+1}^o, consists of the savings plus the interest received and any pensions (b) received. σ^h is the degree of myopia of the individual in forecasting the expected level of pensions. Following Feldstein, $\sigma \in [0,1]$.

2.2.2 Firms

We assume that economy-wide production is determined by a Cobb Douglas of the form:

$$Y_t = (A_t L_t)^{1-\alpha} K_t^{\alpha} \tag{6}$$

where α is the share of capital in output. Y_t represents aggregate output of a homogeneous good and this is determined by the amount of labour, L, and capital, K, available at time t. For ease of manipulation, we assume that technology, A_t, is labour enhancing. We abstract from the impact of social security on the labour supply decision and instead assume that all agents supply one unit of labour inelastically. The economy is endowed with an initial capital stock $K_0 > 0$ and capital depreciates fully from one period to the next. Dividing the production function by $A_t L_t$; in intensive form the production function is given by:

$$y = k^{\alpha} \tag{7}$$

where k is the unit of capital per effective unit of labour. The production function satisfies the usual conditions such that $f(0) = 0$, $f'(k) > 0$, $f''(k) <$

0 and the Inada conditions: $\lim_{k\to 0} f'(k) = \infty$ and $\lim_{k\to\infty} f'(k) = 0$.

Profit maximising perfectly competitive firms pay labour and capital their respective marginal products which for the Cobb-Douglas function is given by:

$$w = (1-\alpha)\,k^{\alpha} \tag{8}$$

$$R = \alpha k^{\alpha-1} \tag{9}$$

where $R \equiv (1+r)$ is the gross rate of interest. By endogenising $w(k)$ and $R(k)$, we now depart from Feldstein who had no capital in his model.

2.2.3 Planner

The role of the planner at a given point in time is restricted to that of maximising the welfare of all individuals living at that point in time. To achieve its objective, the planner operates a Pay-As-You-Go (PAYG) pensions scheme which is in place to ensure that myopic individuals have some income on which they can rely on in their old age. The PAYG scheme operates by taxing those currently active and transferring it to those currently living in retirement. To finance the PAYG scheme, the planner imposes a proportional tax θ on the wage w of the workers. However, though there is a need to protect people who fail to save for their retirement because of myopia, the knowledge that the planner will "bail-out" the myopics out introduces a distortion in the behaviour of the economic agents. As such, the optimal level of pensions will balance the need for protection whilst mitigating the economic costs (Feldstein, 1985). The aggregate amount of tax raised by the planner can thus be expressed as:

$$T_t = \theta_t w_t \left(A_t L_t\right) \tag{10}$$

and this is then redistributed as benefit b_t to those in retirement. In aggregate form this is:

$$B_t = b_t \left(A_{t-1} L_{t-1}\right) \tag{11}$$

Given $A_t L_t = (1+g)(1+n) A_{t-1} L_{t-1}$, the benefit received by each agent in retirement is:

$$b_t = (1+\gamma)\,\theta_t w_t \tag{12}$$

where $(1+\gamma) \equiv (1+n)(1+g)$. It can be deduced that the level of pensions received by the retirees is a function of the wage, the tax rate and the rate of population growth and technological improvement.

2.3 Competitive Equilibrium

Given the households' and the firms' objectives, a competitive equilibrium for the economy can be defined as a sequence of consumption $\{c_t^y, c_t^o\}_{t=0}^{\infty}$ such that:

1. A given sequence of taxes and transfers, $\{\tau w_t, b_t\}_{t=0}^{\infty}$, and the prevailing competitive wages, w_t, and interest rate, R_t, solves the individual's optimisation problem subject to satisfying the Euler equation;

2. Factors of production are paid their marginal products ($w_t = (1-\alpha)k_t^{\alpha}; R_t = \alpha k_t^{\alpha-1}$) and labour and capital markets clear such that $L_t^D = L_t$ and $S_t = K_{t+1}$;

   With complete depreciation[3], $S_t = K_{t+1}$ is a standard condition suggesting that the capital stock in a given period is the savings of the elderly from the previous period.;

[3] We are not referring to depreciation in the normal sense: rather, the old consume capital and the young replenish it.

3. The planner's budget is always balanced hence taxes raised is distributed as benefits in the same period $T_t = B_t$;

4. The economy's resource constraint is always satisfied. In intensive form, the constraint which is defined as the allocation of current output, y_t,

$$y_t = c_t^y + \frac{c_t^o}{(1+\gamma)} + (1+\gamma)k_{t+1} \tag{13}$$

The resource constraint suggests that output at any time is divided between consumption and capital formation. Consumption consists of that of the young and the old.

(1)-(4) define the competitive equilibrium. We can now write the intertemporal budget constraint (IBC) of the agent. The IBC suggests the lifetime consumption of the agents equals their income.

$$c_t^y + \frac{c_{t+1}^o}{R_{t+1}} = w_t(1-\theta) + \frac{\sigma^h b_{t+1}}{R_{t+1}} \tag{14}$$

The Lagrangian and the first order conditions can now be expressed as:

$$\underset{\{c_t^y, c_{t+1}^o\}}{Max} : \ell = \ln c_t^y + \lambda^h \ln c_{t+1}^o - \varrho[c_{t+1}^o - R_{t+1}\{w_t^y(1-\theta) - c_t^y\} + \sigma^h b_{t+1}] \tag{15}$$

$$\frac{\partial \ell}{\partial c_t^y} \quad : \quad \varrho R_{t+1} c_t^y = 1 \tag{16}$$
$$\frac{\partial \ell}{\partial c_{t+1}^o} \quad : \quad \lambda c_{t+1}^o = \varrho \tag{17}$$

where ϱ represents the Lagrangian multiplier. Combining the two first order conditions leads us to the Euler equation, which is the optimal allocation of consumption during the two life periods of the agents:

$$c_{t+1}^o = \lambda^h R_{t+1} c_t^y \tag{18}$$

Based on the Euler equation, the optimal level of consumption and savings of the agents can be expressed as:

$$c_t^{y,h} = \left(\frac{1}{1+\lambda^h}\left[w_t(1-\theta) + \frac{\sigma^h b_{t+1}}{R_{t+1}}\right]\right) \tag{19}$$
$$c_{t+1}^{o,h} = \left(\frac{\lambda^h R_{t+1}}{1+\lambda^h}\left[w_t(1-\theta) + \frac{\sigma^h b_{t+1}}{R_{t+1}}\right]\right) \tag{20}$$
$$s_t^{y,h} = \left(\frac{1}{1+\lambda^h}\left[\lambda^h w_t(1-\theta) - \frac{\sigma^h b_{t+1}}{R_{t+1}}\right]\right) \tag{21}$$

Eqns. (19)-(20) refer to the optimal level of consumption of the young and old during the two periods of their lifetime. This is based on their income and pensions transfers. If the agents have perfect foresight such that $\lambda = \sigma = 1$, then the agents consume half of their income in each period of their lifetime. In the limit that $\lambda \to 0$ and $\sigma = 1$, most of the consumption takes place in the first period. One of the criticisms levelled against a PAYG pensions system is that it can potentially crowd out (private) savings. This can be seen in Eqn. (21) where savings is lower by the extent of the discounted value of pensions. However, the fact that σ relates to the myopia pertaining to the expected level of pensions implies that if $\sigma < 1$, the crowding out is not complete.

Based on definition 2 of the competitive equilibrium and the other condition that capital is crucial for production in that $f(0) = 0$ (and $\lim_{k \to 0} f'(k) = \infty$), we discard Feldstein's notion about the entire population suffering from complete myopia. Instead we assume that myopia is (at worst) partial such that $\lambda > 0$. This is important to ensure there is capital formation in the economy from one period to the next.

Let us assume that the population consists of two types of agents which can be classified according to the level of their myopia, low$\left(\lambda^l\right)$ or high $\left(\lambda^h\right)$ where $1 \geqslant \lambda^l > \lambda^h > 0$. We assume the agents with a low myopia (life-cyclers) make up a proportion μ of the economy and the agents with a high myopia (myopes) make up the remaining $(1 - \mu)$. Capital in the economy is thus made up of the savings of the young born in the previous period. With complete depreciation from one period to the next, capital formation can be described by the following[4]:

$$(1+\gamma)k_{t+1} = \mu s_t^l + (1-\mu)s_t^h \tag{22}$$

We can immediately infer that if all the agents have a low myopia ($\mu = 1$) then, $(1+\gamma)k_{t+1} = s_t^l$. On the other hand, if all the agents suffer from a high degree of myopia ($\mu = 0$), $(1+\gamma)k_{t+1} = s_t^h$. Since $s_t^l > s_t^h$, the two equations imply that the capital stock available to an economy is higher the lower the degree of myopia. Given the above considerations, the steady state capital without pensions (k_0) and with pensions (k_1) can be described by:

$$k_0 = \left(\frac{(1-\alpha)}{(1+\gamma)}\left[\left(\frac{\mu\lambda^l}{1+\lambda^l}\right) + \left(\frac{(1-\mu)\lambda^h}{1+\lambda^h}\right)\right]\right)^{\frac{1}{(1-\alpha)}}$$

$$k_1 = \left(\left(\frac{(1-\theta)(1-\alpha)}{(1+\gamma)}\right)\left[\frac{\alpha\left[\mu\left(1+\lambda^h\right)\lambda^l + (1-\mu)\left(1+\lambda^l\right)\lambda^h\right]}{\alpha\left(1+\lambda^l\right)\left(1+\lambda^h\right) + \theta(1-\alpha)\left[\mu\left(1+\lambda^h\right)\sigma^l + (1-\mu)\left(1+\lambda^l\right)\sigma^h\right]}\right]\right)^{\frac{1}{(1-\alpha)}}$$

We can find that k_0 depends on α, γ, μ and λ whilst k_1 depends on two additional parameters θ and σ. Let us simplify the analysis such that $\lambda^l = \lambda^h = \lambda$ and $\sigma^l = \sigma^h = \sigma$. In that case the steady-state capital is given as:

[4] This is part 2 of the definition of competitive equilibrium written in intensive form.

$$k_0 = \left(\frac{\lambda(1-\alpha)}{(1+\gamma)(1+\lambda)}\right)^{\frac{1}{(1-\alpha)}} \tag{23}$$

$$k_1 = \left(\left[\frac{\lambda\alpha(1-\alpha)(1-\theta)}{(1+\gamma)\left[\alpha(1+\lambda)+\theta(1-\alpha)\sigma\right]}\right]\right)^{\frac{1}{(1-\alpha)}} \tag{24}$$

Proposition 1 *(a) Capital is higher under FF than PAYG*

(b) An increase in λ (i.e, a reduction in myopia), increases k

(c) An increase in σ (i.e., less myopia on expected pensions) reduces k

(d) An increase in θ (i.e, higher PAYG pensions) reduces k

(e) An increase in γ (e.g, higher population growth rate) reduces k

Proof. The Proof is in Appendix 2.1. ■

2.4 Planner's Problem

Given the competitive equilibrium, the planner chooses θ to maximise the "true" welfare function $\{c_t^y, c_t^o, k_{t+1}\}_{t=0}^{\infty}$ subject to the allocation $\{y_t\}_{t=0}^{\infty}$. All agents are given an equal weight. We assume that the utility function remains the same across generations. In steady state, the planner's problem can thus be expressed as maximising the welfare of the young and old, with varying levels of myopia simultaneously. The welfare function can be written as:

$$
\begin{aligned}
V &= (1+\gamma)\left[\mu \ln(c^{y,l}) + (1-\mu)\ln(c^{y,h})\right] \\
&\quad + \left[\mu \ln(c^{o,l}) + (1-\mu)\ln(c^{o,h})\right]
\end{aligned} \tag{25}
$$

The welfare function has two parts: the first part refers to the young whilst the second refers to the old. With population growth and labour-augmenting technological growth, there are $(1+\gamma)$ more young than elderly. The welfare function also reflects the element of heterogeneity in the form of one group having a higher level of myopia than the other.

The current approach could be considered as another departure from Feldstein. Whilst Feldstein simply considered the optimal level of social security, by not considering the full general equilibrium effects and comparing the results with the competitive equilibrium, his results could not suggest whether it was optimal to have a social security scheme at all times. Indeed, there might be circumstances, when it might be optimal to have no social security benefits. People are then responsible for the funding of their own retirement. So, we now aim to find out the optimal level of social security in a general setting before considering specific conditions pertaining to myopia and heterogeneity. For ease of notation, let us denote V_0 as the welfare function under competitive equilibrium (without pensions) and V_1 with PAYG pensions. The decision rule is straight forward:

- If $V_0 > V_1$: It is optimal to have no old age pensions. The pensions system is then FF;
- If $V_0 = V_1$: Having or not having old age pensions does not matter;
- If $V_1 > V_0$: Having old age pensions is optimal. The pensions system is a PAYG scheme.

Case 1: $\lambda^l = \lambda^h = \lambda; \sigma^l = \sigma^h = \sigma$

The special case that all the agents have the same level of myopia eliminates the heterogeneity in the model. This means that the welfare function is simply the sum of the utility of the young and the old: $V = (1+\gamma)\ln(c^y) + \ln(c^o)$

$$V_0 = (1+\gamma)\ln\left[\frac{w_t}{1+\lambda}\right] + \ln\left[\frac{\lambda R_t w_{t-1}}{1+\lambda}\right] \tag{26}$$

$$\begin{aligned} V_1 = {} & (1+\gamma)\ln\left[\frac{1}{1+\lambda}\left(w_t(1-\theta) + \frac{\sigma b_{t+1}}{R_{t+1}}\right)\right] \\ & + \ln\left[\frac{\lambda R_t}{1+\lambda}\left(w_{t-1}(1-\theta) + \frac{\sigma b_t}{R_t}\right)\right] \end{aligned} \tag{27}$$

where $b_t = (1+\gamma)\,\theta w_t$.

We now consider a simplifying case to derive some analytical results. We assume $\sigma = 0$ such that people do not reduce their savings as a result of the provision of pensions.

Case 1a: $\sigma \approx 0$

If the agents do not reduce their savings as a result of pensions, then eqn. (24) reduces to $k_1^0 = \left(\frac{\lambda(1-\alpha)(1-\theta)}{(1+\lambda)(1+\gamma)}\right)^{\frac{1}{(1-\alpha)}}$. Hence, the young consume $\frac{w_t(1-\theta)}{1+\lambda}$. Whilst the elderly planned to consume $\frac{\lambda R_t w_{t-1}(1-\theta)}{1+\lambda}$, they end up consuming an extra b_t in the form of pensions. The "real" function of the planner can now be written as:

$$V_1^0 = (1+\gamma)\ln\left[\frac{(w_t(1-\theta))}{1+\lambda}\right] + \ln\left[\frac{\lambda R_t\,(w_{t-1}(1-\theta))}{1+\lambda} + b_t\right] \tag{28}$$

where the first part refers to the consumption of the young and the second part to the consumption of the old including the pensions. We can now proceed to find the optimal level of tax by setting $\frac{\partial V_1^0}{\partial \theta} = 0$. Assuming $\gamma = 0$, this yields:

$$\theta = \frac{[\lambda - (\alpha\,(\alpha+1))]}{\lambda\,[1 + (1-\alpha^2)]} \tag{29}$$

For the simplifying case, we can find that the optimal level of tax is a function of λ and α. We find that for $\theta > 0$ we require $\lambda > \alpha(\alpha + 1)$. If the level of myopia increases beyond a certain level such that $\lambda < \alpha(\alpha + 1)$, then the feasible optimal level of social security is $\theta = 0$. Hence, at high levels of myopia, it is optimal to have no PAYG pensions and the FF scheme prevails.

Proposition 2 *The higher the myopia on pensions, the lower the optimal PAYG.*

This is confirmed by $\frac{\partial\theta}{\partial\lambda} = \frac{\alpha(1+\alpha)}{\lambda^2[1+(1-\alpha^2)]} > 0$ suggesting that as λ falls, that is myopia on pensions increases, the optimal tax rate falls. This suggests that if there is minimum dissaving as a result of the provision of pensions, then the consumption smoothing role of the PAYG is no longer as important. We have to note that this result is being derived under a very strong assumption whereby the PAYG system is causing minimal disruption to capital formation.

Case 1b: General $\sigma \neq 0$

One of the main criticisms that has been advanced against PAYG schemes is that it acts as a disincentive to save. As such, $\sigma > 0$ might be a more sensible approach. We can then rewrite eqn. (27) as V_1^{σ} where:

$$
\begin{aligned}
V_1^{\sigma} = {} & (2+\gamma)\ln\left[\alpha+(1-\alpha)\sigma\theta\right]+\frac{\alpha+(1+\gamma)}{(1-\alpha)}\ln(1-\theta) \\
& -\left[\frac{\alpha+(1+\gamma)}{(1-\alpha)}\right]\ln\left[\alpha\left(1+\lambda\right)+\theta(1-\alpha)\sigma\right]+Z_1
\end{aligned}
\tag{30}
$$

where Z_1 is a set of parameters independent of the policy term. We find the optimal tax by setting $\frac{\partial V_1^{\sigma}}{\partial \theta}=0$ and this yields:

$$
\frac{(2+\gamma)\left(1-\alpha\right)\sigma}{\left[\alpha+(1-\alpha)\sigma\theta\right]}-\frac{\alpha+(1+\gamma)}{(1-\alpha)}\left[\frac{1}{(1-\theta)}+\frac{\theta(1-\alpha)}{\left[\alpha\left(1+\lambda\right)+\theta(1-\alpha)\sigma\right]}\right]=0
\tag{31}
$$

Let us assume $\alpha=0.25$, $\gamma=0$ and $\sigma=1$ such that there is maximum crowding out from the provision of pensions. If $\lambda=0$, that is the agents are completely myopic, then $\theta^*=0.299$. On the other hand, if the agents are "life-cyclers" and $\lambda=1$, then $\theta^*=0.179$. Thus, there is a need for a lower tax rate if the agents are life-cyclers. This is in line with Feldstein who suggests that complete myopia sets the upper bound on the optimal tax (and benefit) levels.

Impact of myopia on welfare Let us consider eqn. (26) for the FF pensions system. For $\gamma=0$:

$$\frac{\partial V_0}{\partial \lambda} = \left[\frac{(3\alpha - 1)}{(1-\alpha)\lambda}\right] - \left[\frac{(1+\alpha)}{(1-\alpha)\,(1+\lambda)}\right]$$

Hence, an increase in myopia results in a fall in welfare if $\frac{\partial V_0}{\partial \lambda} < 0$. This requires $\left[\frac{3\alpha-1}{1+\alpha}\right] < \left[\frac{\lambda}{1+\lambda}\right]$. This condition is satisfied for all $\alpha < \frac{1}{3}$. Hereunder we show the critical values of α that ensure an increase in myopia leads to a fall in welfare.

λ	1	0.75	0.5	0.25	0
$\alpha <$	$\frac{3}{5}$	$\frac{5}{9}$	$\frac{1}{2}$	$\frac{3}{7}$	$\frac{1}{3}$

Critical Value of α for increasing myopia to lower welfare

2.4.1 Optimal Pensions with Heterogeneity

In the previous section, we have assumed that all individuals have the same discount factor. We now consider a situation of "total" heterogeneity, that is, where the agents vary in terms of their myopia $\left(\lambda^l > \lambda^h\right)$ and the anticipated level of pensions $\left(\sigma^l \neq \sigma^h \neq 0\right)$. The welfare function, V, can be expressed as hereunder with μ reflecting the heterogeneity in terms of the proportion of myopes:

$$
\begin{aligned}
V &= (1+\gamma)\left[\mu \ln(c^{y,l}) + (1-\mu)\ln(c^{y,h})\right] \\
&\quad + \left[\mu \ln(c^{o,l}) + (1-\mu)\ln(c^{o,h})\right]
\end{aligned}
\tag{32}
$$

The steady-state capital follows from eqn. (24) and can be expressed as $k_1 = \left(\frac{a(1-\theta)}{b+c\theta}\right)^{\frac{1}{(1-\alpha)}}$. To find the optimal tax rate, we set $\frac{\partial V}{\partial \theta} = 0$ and this results in:

$$
\begin{aligned}
\frac{\partial V}{\partial \theta} &= (2+\gamma)\left[\frac{\mu\left[\alpha c + a\sigma^l(1+\gamma)\right]}{\left[\alpha(b+c\theta) + a\sigma^l(1+\gamma)\theta\right]} + \frac{(1-\mu)\left[\alpha c + a\sigma^h(1+\gamma)\right]}{\left[\alpha(b+c\theta) + a\sigma^h(1+\gamma)\theta\right]}\right] \\
&\quad + \frac{(3+\gamma)-(5+2\gamma)\alpha}{(1-\alpha)}\left[\frac{1}{(1-\theta)} + \frac{c}{(b+c\theta)}\right] = 0
\end{aligned}
\tag{33}
$$

where a, b and c are non-policy parameters[5]. Given the nature of eqn. (33), we have recourse to simulations to find the optimal tax rate under a general setting. This is done in the next section.

[5] $a = \alpha(1-\alpha)\left[\mu\left(1+\lambda^h\right)\lambda^l + (1-\mu)\left(1+\lambda^l\right)\lambda^h\right]$
$b = \alpha\left(1+\lambda^l\right)\left(1+\lambda^h\right)(1+\gamma)$
$c = (1-\alpha)(1+\gamma)\left[\mu\left(1+\lambda^h\right)\sigma^l + (1-\mu)\left(1+\lambda^l\right)\sigma^h\right]$

2.4.2 Simulations Results and Sensitivity Analysis

Since within a general equilibrium framework the model loses analytical tractability fairly quickly, we resort to simulations to consider the various implications of myopia for the provision of social security. It might be appropriate to highlight at the outset that we would like to consider the simulation results more in terms of giving an indication of the various effects of the existence of social security on some of the key macro economic aggregates of interest. As such, it would be more appropriate to consider the results as giving an indication of the direction in which the variables are moving, rather than the exact magnitude thereof.

The main parameters of interest are λ^h, which determines the level of myopia in the economy; σ^h, which determines the response of savings to the provision of old age pensions; and μ, which determines the proportions of agents with "perfect foresight" and "myopic" agents in the economy. We consider two situations, one with pensions and another without pensions. The one with pensions will be a PAYG scheme, whereas without pensions, there would be a FF scheme in place. We are thus able to compare which of the two schemes yields the highest welfare.

Parameter values The main value for the purpose of these simulation is α, which we set to 0.35. This is in line with the range of parameter values for the share of capital in output-and is driven mainly by the consideration to avoid a situation of dynamic inefficiency for the case when all the agents have foresight. The latter yields the lowest interest rate and with values of $\alpha < 0.35$ could potentially be dynamically inefficient such that $r < \gamma$. We assume the population growth rate to be 1% whilst the rate of technological progress is set at 2.5%. Hence, $\gamma = 3.53\%$.

Simulation Set 1$\left(Baseline : \lambda^l = 1; \lambda^h \to 0; \sigma = 1\right)$: We initially set all agents to have perfect foresight and then allow one proportion to suffer from myopia. We allow for the degree of myopia to increase. We intially set the proportion of myopes to $\mu = 0.25$ but also compare with values of $\mu = 0.5$ and 0.75, i.e., allowing a greater proportion of life-cyclers in the economy. In the first set of results, we assume that all agents reduce their savings by the full amount of the pensions. We compare the results from the PAYG scheme with the FF scheme to consider which one yields the highest welfare.

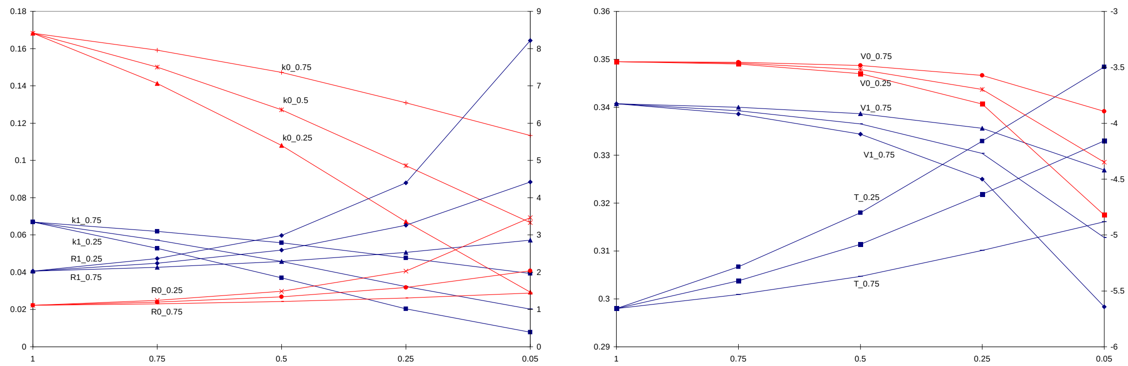


Figure 1: Impact of Increasing Myopia (Sim 1)

Impact on capital and interest rate: The preliminary results in so far as capital is concerned still hold. For any level of myopia, steady-state capital is higher under the FF scheme than the PAYG scheme. The higher the proportion of myopes in the economy, the lower the capital and hence the higher the interest rate. When the agents have perfect foresight, the level of capital is the same. However, as myopia sets in, as the proportion of myopes increase, there is a wedge between the various capital levels that widens. The result can thus be summed up as follows: the higher the myopia and the proportion of myopes, the lower the capital.

Impact on tax rate and Welfare: With perfect foresight, the tax rate is around 30 percent. As the proportion of myopes increases, the tax rate increases as expected; however the increase is relatively small. For instance, when 25 percent of the agents suffer from a high level of myopia whereby they allocate only a 5 percent weight to the future, the tax rate is around 31.6 percent. Allowing the proportion of myopes to increase to 50 percent and 75 percent leads to increase in the tax rate to around 33.3 and 34.8 percent, respectively. This is consistent with Feldstein who suggests that myopia sets the upper-bound for the tax rate.

Welfare is unambiguously higher at all levels under the FF scheme than under the PAYG scheme. As myopia increases and the proportion of myopes increases, welfare falls. This can be rationalised on account of several factors: firstly, the adverse impact of myopia on capital stock which results in lower output and wages (hence consumption) and higher interest rates. On the other hand, for the PAYG scheme, over and above myopia, $\sigma = 1$ leads to a full reduction in savings and hence a lower capital stock. Moreover, the tax rate increase as well. All these combine to yield a lower welfare.

Simulation Set *2* $\left(\lambda^l = 1; \lambda^h \to 0; \sigma = 0.5\right)$: We now consider a situation whereby the agents do not reduce their savings by the full amount of the pensions. One of the key arguments against PAYG pensions schemes is that it affects capital formation adversely and as a result inhibits growth. In this model, we can control for this distortion through the parameter σ. Instead, we assume that $\sigma = 0.5$, such that they reduce their savings by only half the level of pensions. The experiment is thus similar to experiment 1 *except* for $\sigma = 0.5$.

Impact on capital and interest rate: The first finding suggests that none of the values under the FF scheme change compared to the baseline. This can be rationalised on the account that the σ parameter only relates to the distortion to savings resulting from pensions. Since this parameter does not affect the FF scheme, the values for the FF scheme do not change.

As in the baseline, the capital stock is lower under the PAYG system as opposed to the FF scheme resulting in lower output and wages and higher interest rates. However, the resulting fall in capital stock is now only marginally lower compared to the FF scheme and as a result compared to the baseline PAYG, output and wages are higher and the interest rate is lower. The same dynamics as previously apply in that as the level of myopia and the proportion of myopes increase, capital falls.

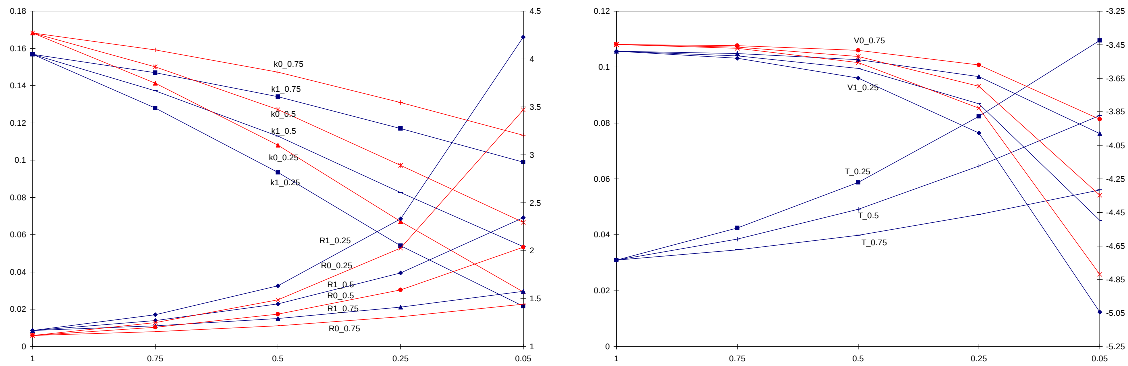


Figure 2: Impact of Increasing Myopia (inc. pensions) (Sim 2)

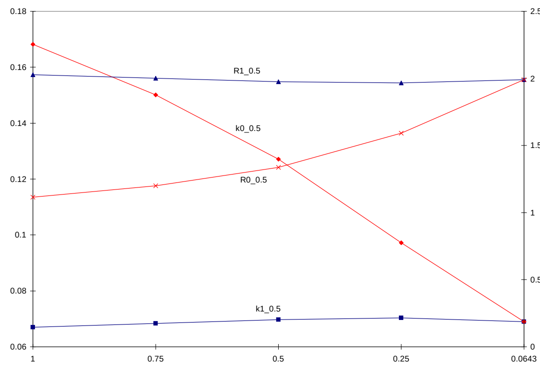
Impact on tax rate and welfare: The tax rate falls considerably. For the case when all the agents have perfect foresight, the tax rate stands at around 3.09 percent only. As the proportion of myopes increases and the level of myopia increases, the tax rate increases. When 25 percent of the agents suffer from near full myopia, the tax rate is around 5.6 percent. When 75 percent suffer from the same level, the tax rate is around 11 percent. Hence, we find that when the tax rate is lower when individuals save a higher proportion of their income when working - and the result is consistent with consumption smoothing in the sense that if the agents are already saving, then the importance of pensions falls. The results remain consistent with Feldstein in that the higher the level of myopia, the higher the tax rate.

Welfare is unchanged under the FF scheme but is still higher than the

PAYG scheme. Compared to the baseline, welfare under the PAYG is now higher at all levels. The same findings hold in that as the level of myopia and the proportion of myopes increase, welfare falls.

Simulation Set ***3*** $\left(\lambda^l = 1, \sigma^l = 1; \lambda^h = \sigma^h \to 0; \mu = 0.5\right)$: We now consider a case where the agents with perfect foresight make their decisions incorporating the full amount of their pensions (hence reducing their savings by an equal amount). On the other hand, the proportion of myopes reduce their savings by the same extent as their myopia. Hence, as their myopia increases, they expect a lower level of pensions and dissave less. We assume that half of the agents are myopes and the other half are lifecyclers ($\mu = 0.5$).

Impact on capital and interest rate: There are two key results that emerge in so far as capital is concerned. Though FF capital remains higher than the PAYG, we note that the FF capital falls as the level of myopia increases thereby resulting in an increase in the interest rate. However, under the PAYG scheme, capital remains more or less constant. This suggests that the fall in σ results in a lower reduction in savings as a result of the provision of pensions and as such capital does not fall, as in the baseline case. We note that the capital under the FF and PAYG schemes are almost the same when



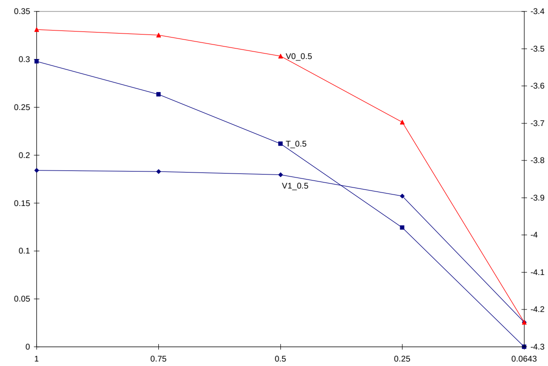
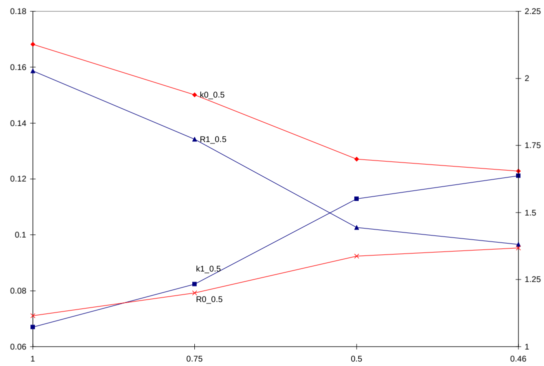


Figure 3: Myopia on both counts (Sim 3)

$\lambda^h = \sigma^h = 0.0643$. This is due to the fact that for $\lambda^h = \sigma^h = 0.0643$, the optimal tax rate is zero. Hence, $k_0 = k_1$.

Impact on tax rate and welfare: The tax rate falls gradually as the level of myopia increases. When all the agents have perfect foresight, the tax rate is around 30 percent. However as myopia sets in, the tax rate falls such that for $\lambda^h = \sigma^h = 0.0643$, the tax rate is zero. Beyond this level a FF scheme is optimal since the feasible optimal tax rate is zero. We also note that in the above simulations, there is a fall in welfare as myopia increases but there is a convergence in welfare under FF and PAYG such that for $\lambda^h = \sigma^h = 0.0643$ the welfare under the two schemes are the same. But this is to be expected since the tax rate at this level is zero.



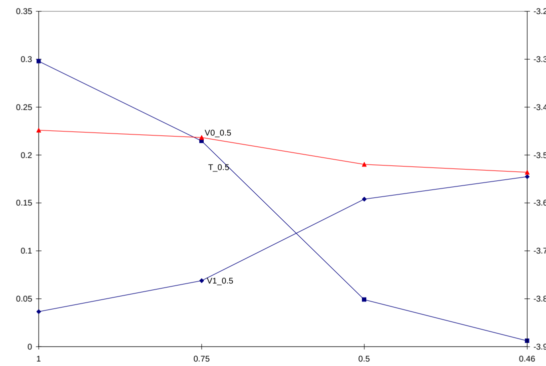


Figure 4: Myopia on Pensions by both Agents (Sim 4)

Simulation Set *4* $\left(\lambda^l = 1, \sigma^l = \sigma^h \rightarrow 0; \lambda^h \rightarrow 0; \mu = 0.5\right)$: We now consider the case where the agents with perfect foresight face a level of myopia as regard to pensions which is similar to that faced by the myopes. As a result, they do not reduce their savings by the full amount of the pensions.

Impact on capital stock,output, wages and interest rate Capital under the PAYG scheme is now higher than in the baseline. We note a convergence in capital such that FF capital falls as myopia increases whilst it increases in the case of PAYG since as σ falls, the dissaving from pensions falls as well. This explains the convergence in the interest rates as well.

Impact on tax rate and welfare The tax rate falls rather sharply such that beyond $\sigma = \lambda^h = 0.46$, the optimal tax is zero. There is also a convergence in welfare but FF dominates. Whilst welfare falls in the FF case as myopia increases, it increases in the case of PAYG. In part this is due to the increasing capital (and falling tax rate). Welfare is the same around $\sigma = \lambda^h = 0.46$ but at this level, the tax rate is nearly zero.

2.5 Conclusions

With the decline in the fertility rates combined with an increase in longevity, the long term sustainability of the PAYG system has come under threat. Indeed, various countries have undertaken a reform of their pensions system or are considering ways to deal with an ageing population. In this paper we have extended the findings of Feldstein whereby the full general equilibrium effects of the PAYG pensions are captured. We depart from Feldstein in two key ways: firstly, we allow for capital formation and the distortionary cost of capital formation on it and secondly we also compare our findings with a FF scheme. Thus, we are able to see which of the FF or PAYG yields a higher welfare. Our main results in so far as the PAYG is concerned remain consistent with Feldstein. We find that the PAYG leads to a reduction in capital and that the higher the myopia, the higher the tax rate. However, if the myopia also affects the agent's ability to anticipate the exact amount of pensions he receives, then the crowding out with the PAYG is not complete. In such cases, the less the crowding out, the lower the tax rate. We also find that there is a range of parameters for which it is not optimal to have a social security system in place for consumption smoothing. Under such circumstances, a FF system will operate whereby all the agents are responsible for the provision of their own pensions. Consistent with Aaron (1966), we also find that in a dynamically efficient economy, a FF system yields a higher welfare than a PAYG.

Appendix 2.1

Proof for proposition 1

(1a): The condition for $k_0 > k_1$ $[Eqn.(23) - (24)]$ reduces to:

$\alpha\,(1+\lambda) + (1-\alpha)\sigma > 0.$

This is always satisfied.

(1b): $\frac{\partial k_0}{\partial \lambda} = \frac{1}{(1-\alpha)}\left[\frac{(1-\alpha)}{(1+\gamma)(1+\lambda)^2}\right]^{\frac{\alpha}{(1-\alpha)}} > 0$

(1c): $\frac{\partial k_1}{\partial \sigma} = \frac{1}{(1-\alpha)}\left[\frac{-(1-\alpha)^2\lambda\alpha\theta(1-\theta)}{(1+\gamma)[\alpha(1+\lambda)+\theta(1-\alpha)\sigma]^2}\right]^{\frac{\alpha}{(1-\alpha)}} < 0$

(1d): $\frac{\partial k_1}{\partial \theta} = \frac{1}{(1-\alpha)}\left[\frac{-\lambda\alpha(1-\alpha)(1-\theta)[\alpha(1+\lambda)+(1-\alpha)\sigma]}{(1+\gamma)[\alpha(1+\lambda)+\theta(1-\alpha)\sigma]^2}\right]^{\frac{\alpha}{(1-\alpha)}} < 0$

(1e): $\frac{\partial k_1}{\partial \gamma} = \frac{1}{(1-\alpha)}\left[\frac{-\lambda\alpha(1-\alpha)(1-\theta)[\alpha(1+\lambda)+\theta(1-\alpha)\sigma]}{(1+\gamma)[\alpha(1+\lambda)+\theta(1-\alpha)\sigma]^2}\right]^{\frac{\alpha}{(1-\alpha)}} < 0$

3 PENSIONS OR INCOME SUPPORT: WHICH IS THE OPTIMAL REDISTRIBUTION INSTRUMENT?

3.1 Introduction

The role of governments in economies has been widely debated by the various schools of thought. However, there is some agreement that there are some basic functions that a government needs to undertake to ensure a smooth running of an economy. Whilst the need to address pervasive market failures in terms of externalities and inefficiencies are among the most important, redistribution has also been at the fore of the policy debate. The need for redistribution can be traced back at least to Adam Smith (1776) who suggested, in The Wealth of Nations, "*No society can surely be flourishing and happy, of which the far greater part of the members are poor and miserable*".

In this paper we consider what is the optimal way to redistribute from high productivity individuals (the rich) to low productivity individuals (the poor). We aim to investigate whether redistribution should be in the form of income support, that is given to people when they are young or in the form of a pay-as-you-go (PAYG) pensions. Although the term PAYG is used, in a somewhat general sense, the pensions system is in fact "means tested" where only the poor receive pensions. Pensions thus provide some form of social insurance and help alleviate poverty and/or reduce inequality.

To answer the questions posed, we develop a two-period overlapping generations (OLG) general equilibrium model for a steady state economy with endogenous capital formation to take account of some of the distortionary elements of redistribution. Following Galor (1992), the steady state can be considered as the representative framework within which infinitely many generations evolve. Our model also provides for a richer analytical framework, in that we do not restrict our work to the existence of an optimal level of transfer in a PAYG system, but also consider the instances where a fully funded system yields a higher level of welfare. Crucial to our redistribution argument is an element of heterogeneity in the form of productivity of the poor relative to the rich. In this setting, the rich fund for their own pensions and they pay a constant proportion of their wages to the planner in the form of taxes. The planner is then faced with the problem of deciding how best to allocate a non-negative transfer to the poor. We assume this can take the form of either income support or alternatively a pay-as-you-go (PAYG) pensions scheme[6]. By definition, if the means of redistribution is taken to be income support, then the poor have to save for their own retirement and the system would be fully funded (FF). On the other hand, if the redistribution is done through pensions, then the system would be a hybrid with the pensions of the rich operating under a FF scheme whilst that of the poor would be under a PAYG scheme. Our paper thus brings together two approaches that have usually been considered on an individual basis.

We are able to derive some key analytical results. These can be summarised as:

1. Redistribution imposes costs on the economy in terms of capital formation. Consistent with Feldstein (1974), a PAYG system leads to a lower capital stock. This is because the beneficiaries (the poor) reduce their savings in anticipation of pensions whilst the rich have a lower level of savings because of a lower disposable income;

2. Capital is always higher under a FF system than the PAYG. Moreover,

[6]Though we only consider the transfer in either one form or another, it is also possible for the planner to provide the income to the poor during the two periods of their lives. A poor agent would thus receive income support when young and pensions when old. This is considered in Chapter 4.

if all the agents have the same discount factor, then intragenerational redistribution does not affect capital formation. On the other hand, intragenerational redistribution can adversely affect capital formation when the poor discount the future more than the rich;

3. An increase in the proportion of the rich and/or productivity reduces the tax rate for intragenerational transfers;

4. The utility of the rich falls as the tax rate rises whilst that of the poor increases;

5. An increase in the proportion of the rich and/or productivity increases welfare unambiguously;

6. An increase in the population growth rate reduces welfare through a decrease in capital.

Given the nature of the OLG model, we lose analytical tractability beyond a certain point and have recourse to simulations to show some of the general equilibrium effects. The parameter values are in line with the macroeconomic literature and are chosen to yield a dynamically efficient economy in the baseline case. There are parameter ranges for which it is optimal to have no redistribution through pensions. The results also show that for some parameters the economy can become dynamically inefficient $(r < n)$. When this is the case, there has been an overaccumulation of capital and the prevailing interest rate is less than the population growth rate. According to Samuelson (1975), welfare can be increased by reducing capital in such instances. As such, a PAYG scheme can be used to reduce capital and restore dynamic efficiency.

Some of the simulation results can be summarised as follows:

1. Although redistribution introduces some elements of distortion in the economy, we find that intragenerational transfers always leads to an increase in welfare compared to the initial situation without redistribution. This can be rationalised on account of the fact that the welfare gains of the poor at least exceeds the welfare loss of the rich;

2. Though transfers through pensions are also welfare enhancing, there exists a set of parameters for which it is optimal to have no redistribution through pensions;

3. The Aaron (1966) condition holds in the case of dynamic inefficiency $(r < n)$. If the economy is dynamically efficient, redistributing through pensions yields higher welfare;

4. When the economy is dynamically efficient $(r > n)$, income support prevails. However, as the economy approaches dynamic inefficiency, there is a small range when with $r \gtrapprox n$ for which it is optimal to redistribute through pensions. We suggest the PAYG pensions ensures the economy remains dynamically efficient;

5. Intragenerational transfers are optimal when an economy is characterised by a small proportion of rich and the poor have a low productivity since such an economy will be dynamically efficient. Otherwise, in an economy where there is no wide differences in the proportion of rich and the productivity of the poor, a PAYG might deliver a higher level of aggregate welfare since such an economy can potentially be dynamically inefficient.

This paper can be related to two strands of the literature, namely: pensions and ageing, and inequality. With the ageing of the population of most countries across the world and the associated implications for the sustainability of PAYG pensions schemes, there have been calls (World Bank, 1994; Gill et al., 2004) to review the way pensions are provided. The current work can be considered as an extension to analyse the implications of putting in place a targeted approach to the provision of pensions. Thus, only the poor receive pensions. Our work could also be seen as a departure from the consumption smoothing role of pensions, as in Feldstein (1985), or cases where pensions is motivated by altruism (Hansson and Stuart, 1989) or that of risk-sharing as prescribed by Shiller (1999), Conesa and Krueger (1999) and Bohn (1999). We consider pensions as social insurance.

There is a wide array of literature that has focused, directly or indirectly, on inequality and the various channels through which it affects growth and welfare since Kuznets (1955). Since economies are seldom homogeneous, with agents differing across skills and asset endowments to highlight but two, there is a growing belief that redistribution can have major politico-economic ramifications and implications for policy. Redistribution is often considered as one of the routes through which social justice and efficiency can be promoted by reducing inequality and supporting those at the lower end of the economy. Besides the philanthropic arguments, there have been growing concerns that inequality can be harmful for growth and too skewed a distribution of assets and income can have damaging consequences (Persson and Tabellini (1994); Alesina and Rodrik (1994); Alesina and Perrotti (1996)). Intragenerational transfer is the instrument *par excellence* to bridge the gap between the rich and the poor but there is a limited number of papers that deal explicitly with how intergenerational transfers can be used to reduce inequality. Haveman (1988) argues that intragenerational inequality reducing transfer works best since it increases the opportunities of the young poor. Krueger and Kubler (2006) look at intergenerational transfers but more from the perspective where markets are missing.

The current work can be seen as an extension to a recent paper by Conde-Ruiz and Galasso (2005) who consider the various aspects of intra and intergenerational redistribution. They find that with sufficient inequality in earnings and elderly in the economy, there is an equilibrium that supports the existence of both intra and intergenerational redistribution. We consider the macroeconomic implications of redistribution where pensions aims to achieve some form of social insurance. Most of the papers that have considered inter and intra generational redistribution simultaneously have done so from a political economy perspective whereby agents with differing abilities get to vote on the level of taxes and benefits (Galasso and Profetta, 2002). As is often the case with those models, the relative positioning of the median voter, or the skewness of the distribution of income, is one of the key drivers of the results (example, Galasso and Profetta, 2004, 2007). Razin *et al.* (2002) consider how the optimal level of tax and social transfers vary with changes in the dependency ratio when people have the option of acquiring education - at a cost. Their theoretical finding is ambiguous because of the conflict arising out of ageing between the young voter, who wants to have lower taxes, and the median voter who is more likely to be poor (based on the mechanics of the model), who will want a higher level of redistribution. Hence, the ambiguity. However, their empirical work suggests the first effect dominates the second and ageing leads to a reduction in the welfare state. Bossi (2007) investigates the political economy considerations of which form of social security is adopted in an economy where one generation is faced with the risk associated with a financial asset such as a downturn in the stock market. His findings suggest the existence of such risks is one of the major explanations for the political prevalence of PAYG systems.

The rest of this paper is as follows: in Section 2 we describe the set up of the economy. Section 3 derives the competitive equilibrium whilst Section 4 considers the planner's problem in terms of finding the optimal level and timing of transfer. Section 5 tests the model to a sample of countries. Section 6 concludes and provides some potential extensions.

3.2 The Economy

Economic activity takes place over infinite discrete time $t \in \{0, 1..., \infty\}$ without uncertainty. The economy consists of two types of utility maximising agents $h(r, p)$, profit maximising firms and a welfare maximising social planner. At each time t, two factors, an amount of capital, k, and labour, l, are available as inputs to production and a homogeneous good is produced.

3.2.1 Agents

At t, two generations live simultaneously - one generation is young and the other is old. Population grows at a constant rate n and therefore, at any time t, there are $(1 + n)$ more (young) workers than (old) retirees. The population at t can thus be expressed as: $L_t + L_{t-1} = (2 + n)L_t$. L_t refers to the agents born at t. Following Samuelson (1958), the composition of the population is considered to be stationary. Therefore, the proportions and types of individuals remain the same across generations.

Each young agent is endowed with one unit of labour which he provides inelastically. The labour endowment in the second period of life is zero. Agents differ according to their productivity $\psi \in (0, 1]$, which in turn determines the wage they receive. There are high productivity agents (the rich,

r) and low productivity agents (the poor, p). To simplify the argument, the productivity of the rich (r) is normalised to 1 and hence any agent with $\psi < 1$ is considered as poor (p). We assume that the rich make up a proportion π of the economy and hence the poor make up for the remaining $(1 - \pi)$.

Agents derive utility solely out of consumption and they are non-altruistic. They are thus born without assets and do not leave bequests. At time t the young agent chooses his level of consumption and savings to maximise utility whilst the old agent lives off his savings (and any transfer).

The inter-temporal optimisation problem can be expressed as maximising $u^h(c_t^h, c_{t+1}^h)$ subject to the budget constraints which vary according to the individual's type and mode of transfer in operation. For ease of manipulation, the intertemporal utility function is taken to be additive and log-linear. The utility function is thus strictly concave, since more consumption is preferred to less, and twice differentiable: $u'(c) > 0$ and $u''(c) < 0$. The function also satisfies $\lim_{c \to 0} u'(c) = \infty$ such that subject to its disposable income, the household will always choose a positive level of consumption when maximising life-cycle utility. The rich agent's problem can be expressed as:

$$\underset{\{c_t^y, c_{t+1}^o, s_t\}}{Max} : U^r = \ln c_t^{y,r} + \beta \ln c_{t+1}^{o,r} \tag{1}$$

subject to:

$$c_t^{y,r} = w_t(1-\tau) - s_t^{y,r} \tag{2}$$

$$c_{t+1}^{o,r} = R_{t+1} s_t^{y,r} \tag{3}$$

Let us consider in the first case, the transfer from the young rich to the young poor. The poor agent's problem will thus be of the following form:

$$\underset{\{c_t^y, c_{t+1}^o, s_t\}}{Max} : U^p = \ln c_t^{y,p} + \theta \ln c_{t+1}^{o,p} \tag{4}$$

The maximisation problem of the two agents are broadly similar, in that they maximise consumption over both periods of their lifetime; except for the subjective discount factor, $\theta \leq \beta \in (0,1)$. The disparity in the discount factor can be rationalised along the lines of Becker (1990) who suggests that the poor tend to discount the future at a higher rate. Over and above the disparity in productivity, with both agents providing one unit of labour inelastically, the budget constraints also differ in that the rich are taxed a constant proportion $\tau \in (0,1)$ out of their income and this is redistributed to the poor. Disposable income is then allocated *à la* Diamond (1965) between present consumption and savings. $R \equiv (1+r)$ is the gross rate of return on savings. Irrespective of the transfer mode, the rich consume only their savings in retirement. In turn, the budget constraints of the poor will depend on the mode of transfer in operation. With redistribution in the first period, their consumption is as follows:

$$c_t^{y,p} = \psi w_t + b_t - s_t^{y,p} \tag{5}$$

$$c_{t+1}^{o,p} = R_{t+1} s_t^{y,p} \tag{6}$$

b_t represents the intragenerational redistribution from the rich to the poor. We will refer to b_t as income support. We find that under this scheme, the poor also have only their savings to rely on when old. The pensions system in this set up thus approximates to a fully-funded scheme where everyone is responsible for the provision of their own consumption in retirement through their savings. On the other hand, if the redistribution takes place in the second period of the lifetime, the consumption of the poor will be:

$$c_t^{y,p} = \psi w_t - s_t^{y,p} \tag{7}$$

$$c_{t+1}^{o,p} = R_{t+1} s_t^{y,p} + p_{t+1} \tag{8}$$

p_{t+1} represents the intergenerational redistribution from the young rich to the old poor. We will refer to p_{t+1} as pensions. With intergenerational redistribution, a hybrid pensions system arises in that the rich fund for their own retirement but the poor have an additional source of income in the form of pensions.

3.2.2 Firms

A large number of identical firms produce a homogeneous good using an identical economy-wide Cobb Douglas production function. The production function can be represented as $Y = K^{\alpha} L^{1-\alpha}$, where α is the share of capital in production.

In intensive form, the production function reduces to: $y = k^{\alpha}$. Firms maximise profit by taking factor prices, which are paid their marginal products in a competitive setting, as given. It is assumed that the labour market clears such that labour demand equals labour supply and the wages received by a worker depends on his level of productivity. The economy is endowed with an initial capital stock $K_0 > 0$ and capital depreciates fully from one period to the next. We assume no technological change. The production function satisfies the usual conditions such that $f(0) = 0$, $f'(k) > 0$, $f''(k) < 0$ and the Inada conditions: $\lim_{k \to 0} f'(k) = \infty$ and $\lim_{k \to \infty} f'(k) = 0$.

3.2.3 Social Planner

The social planner is assumed to be a benevolent welfare maximiser. Its role is limited to that of raising taxes from the currently active rich and transferring it to the poor, young or old, depending on which mode yields the highest welfare level. The total taxes (T) raised can be represented as a proportion of wages as follows: $T_t = \pi L_t \tau w_t$. If this is redistributed to the young (poor), the total benefits (B) they receive is: $B_t = \pi L_t \tau w_t$. If the transfer is done as a PAYG pensions scheme, the poor thus receive: $P_t = \pi L_{t-1} \tau w_t$. With a stationary population and commitment device, a poor agent is thus faced with receiving **either** of the following benefit levels:

$$\text{FF} \quad : \quad b_t = \frac{\pi \tau_1 w_t}{(1-\pi)} \tag{9}$$

$$\text{PAYG} \quad : \quad p_{t+1} = \frac{\pi(1+n)\tau_2 w_{t+1}}{(1-\pi)} \tag{10}$$

Since the redistribution affects the capital accumulation process, we can safely assume that the wages will differ in both settings and the optimal level of tax rate will also be different. Hence we can also infer that the level of benefits, either b or p, would be different under both settings. Some comparative statics reveal that $\partial b/\partial \pi = \tau_1 w_t/(1-\pi)^2 > 0$ suggesting that as π, the proportion of rich increases, the level of benefits b increase as well. For the PAYG system, $\partial p/\partial\pi = (1+n)\,\tau_2 w_{t+1}/(1-\pi)^2 > 0$; $\partial b/\partial n = \pi\tau_2 w_{t+1}/(1-\pi) > 0$, implying that ceteris paribus as the population growth rate increases, the generosity of the pensions scheme increases as well.

3.3 Competitive Equilibrium

Given the households' and the firms' objectives, a competitive equilibrium for the economy can be defined as a sequence of consumption $\{c_t^y, c_t^o\}_{t=0}^{\infty}$ such that:

1. A given a sequence of taxes and transfers, $\{\tau w_t, b_t/p_{t+1}\}_{t=0}^{\infty}$, and the prevailing competitive wages, w_t, and interest rate, R_t, solves the individual's optimisation problem subject to satisfying the Euler equation;

2. Factors of production are paid their marginal products. Hence, $w_t = (1-\alpha)k_t^{\alpha}$ and $R_t = \alpha k_t^{\alpha-1}$. Labour and capital markets clear, hence $L_t^D = L_t$ and $S_t = K_{t+1}$;

3. Irrespective of the mode of transfer, the planner's budget is always balanced. Hence taxes raised are redistributed as benefits in the same period $T_t = B_t$;

4. The economy's resource constraint is always satisfied. In intensive form, the constraint suggests that output at any time is allocated between consumption and capital formation. Consumption consists of that of the young and the old; rich and poor.

$$y_t = c_t^y + \frac{c_t^o}{(1+n)} + (1+n)k_{t+1} \tag{11}$$

Given the above definition of competitive equilibrium, the agent has to choose his level of consumption and savings subject to the budget constraint to maximise utility. The intertemporal budget constraint (IBC) suggests the present value of lifetime consumption equals lifetime disposable income. The IBC of the rich can be expressed as:

$$c_t^{y,r} + \frac{c_{t+1}^{o,r}}{R_{t+1}} = w_t(1-\tau) \tag{12}$$

The IBC of the poor will vary with the mode of transfer - with income support, there is no discounting, whilst any pensions received as an elderly will be discounted. The IBCs under the two settings can thus be written as:

$$\text{FF:} \qquad c_t^{y,p} + \frac{c_{t+1}^{o,p}}{R_{t+1}} = \psi w_t + b_t \tag{13}$$

$$\text{PAYG:} \qquad c_t^{y} + \frac{c_{t+1}^{o}}{R_{t+1}} = \psi w_t + \frac{p_{t+1}}{R_{t+1}} \tag{14}$$

Given the intertemporal budget constraint, we can solve for the competitive equilibrium for the poor when the transfer takes place in the first period. The Lagrangian and the first order conditions are:

$$\underset{\{c_t^y, c_{t+1}^o\}}{Max} : \ell = \ln c_t^{y,p} + \theta \ln c_{t+1}^{o,p} - \lambda[c_{t+1}^{o,p} - R_{t+1}\{w_t^{y,p} + b_t - c_t^{y,p}\}] \tag{15}$$

$$\frac{\partial \ell}{\partial c_t^y} \quad : \quad \lambda R_{t+1} c_t^y = 1 \tag{16}$$

$$\frac{\partial \ell}{\partial c_{t+1}^o} \quad : \quad \theta c_{t+1}^o = \lambda \tag{17}$$

where λ represents the Lagrangian multiplier. Combining the two first order conditions yields the Euler equation, which is the optimal allocation of consumption during the two periods of the agent's lifetime:

$$c_{t+1}^o = \theta R_{t+1} c_t^y \tag{18}$$

Using the Euler equation, the optimal consumption and savings of a poor utility maximising agent who receives income support can be described as:

$$
\begin{aligned}
c_t^{y,p} &= \left(\frac{1}{1+\theta}\left[\psi w_t + b_t\right]\right) && (19)\\
c_{t+1}^{o,p} &= \left(\frac{\theta R_{t+1}}{1+\theta}\left[\psi w_t + b_t\right]\right) && (20)\\
s_t^{y,p} &= \left(\frac{\theta}{1+\theta}\left[\psi w_t + b_t\right]\right) && (21)
\end{aligned}
$$

Using the same approach, the optimal level of consumption and savings of the poor who receive a transfer in the second period of their life in the form of pensions can be expressed as:

$$
\begin{aligned}
c_t^{y,p} &= \left(\frac{1}{1+\theta}\left[\psi w_t + \frac{p_{t+1}}{R_{t+1}}\right]\right) && (22)\\
c_{t+1}^{o,p} &= \left(\frac{\theta R_{t+1}}{1+\theta}\left[\psi w_t + \frac{p_{t+1}}{R_{t+1}}\right]\right) && (23)\\
s_t^{y,p} &= \left(\frac{1}{1+\theta}\left[\theta\psi w_t - \frac{p_{t+1}}{R_{t+1}}\right]\right) && (24)
\end{aligned}
$$

The main difference between the two modes of transfer for the poor is the fact that pensions cause savings to be lower - potentially crowding out capital. We first consider the optimal level of consumption and savings of the rich - which can be expressed in a similar manner under both settings since the rich only pay out taxes and receive no benefits. However, the fundamental difference will arise when the tax rates τ_1 and τ_2 are different under the two settings.

$$
\begin{aligned}
c_t^{y,r} &= \left(\frac{1}{1+\beta}\left[w_t(1-\tau)\right]\right) &(25)\\
c_{t+1}^{o,r} &= \left(\frac{\beta R_{t+1}}{1+\beta}\left[w_t(1-\tau)\right]\right) &(26)\\
s_t^{y,r} &= \left(\frac{\beta}{1+\beta}\left[w_t(1-\tau)\right]\right) &(27)
\end{aligned}
$$

It can be seen that if the agents do not discount the future, that is, $\beta = \theta = 1$, then the agents consume exactly half of their income and save the remaining half. Conversely, if $\beta = \theta \approx 0$, then most of the consumption takes place during the first period of lifetime.

Having derived the optimal levels of consumption and savings, it is now possible to consider capital formation. For comparative purposes, the steady-state capital under three settings are derived: without redistribution (k_0); with redistribution in first period (k_1) and second period (k_2) respectively. In all three cases, with complete depreciation of capital from one period to the next, the capital formation process follows:

$$
(1+n)k_{t+1} = \pi s_t^{y,r} + (1-\pi)s_t^{y,p} \tag{28}
$$

implying that the capital available per worker in the current period is the savings of the rich and the poor carried over from the previous period.

We thus find that the proportions of rich and poor play an important role in capital formation. The steady state capital in the three cases are:

$$k_0 = \left(\frac{(1-\alpha)}{(1+n)} \left[\left(\frac{\beta\pi}{1+\beta} \right) + \left(\frac{\theta(1-\pi)\psi}{1+\theta} \right) \right] \right)^{\frac{1}{(1-\alpha)}} \tag{29}$$

$$k_1 = \left(\frac{(1-\alpha)}{(1+n)} \left[\left(\frac{\beta\pi(1-\tau_1)}{1+\beta} \right) + \left(\frac{\theta\left[(1-\pi)\,\psi + \pi\tau_1\right]}{1+\theta} \right) \right] \right)^{\frac{1}{(1-\alpha)}} \tag{30}$$

$$k_2 = \left(\frac{(1-\alpha)}{(1+n)} \left[\frac{(1+\theta)\,\alpha}{(1+\theta)\,\alpha + \pi\tau_2(1-\alpha)} \right] \left[\left(\frac{\beta\pi(1-\tau_2)}{1+\beta} \right) + \left(\frac{\theta(1-\pi)\psi}{1+\theta} \right) \right] \right)^{\frac{1}{(1-\alpha)}} \tag{31}$$

k_0 represents steady state capital without any redistribution. We can see from eqns. (30) and (31) that k_1 and k_2 depend on the additional parameters τ_1 and τ_2 respectively. To carry out some comparative statics we consider the special case where all the agents have the same discount factor such that $\beta = \theta = \delta$. Then, the steady state capital can be simplified to:

$$k_0 = \left(\frac{\delta(1-\alpha)\left[\pi + (1-\pi)\psi\right]}{(1+n)(1+\delta)} \right)^{\frac{1}{(1-\alpha)}} \tag{32}$$

$$k_1 = \left(\frac{\delta(1-\alpha)\left[\pi + (1-\pi)\psi\right]}{(1+n)(1+\delta)} \right)^{\frac{1}{(1-\alpha)}} \tag{33}$$

$$k_2 = \left(\frac{\delta(1-\alpha)\alpha\left[\pi(1-\tau_2) + (1-\pi)\psi\right]}{(1+n)\left[(1+\delta)\alpha + \pi\tau_2(1-\alpha)\right]} \right)^{\frac{1}{(1-\alpha)}} \tag{34}$$

Proposition 3 *(a) An increase in δ increases k*

(b) An increase in π increases k

(c) An increase in ψ increases k

(d) An increase in n reduces k

(e) When the agents have the same discount factor, intragenerational redistribution does not affect capital.

(f) Intergenerational redistribution always lowers capital formation.

Proof. The Proof is in Appendix 3.1. ■

Some of the basic results can be summarised as: $\frac{\partial k_0}{\partial n} < 0$ implying that as the population growth rate increases, capital available per worker falls; $\frac{\partial k_0}{\partial \delta} = \frac{\partial k_1}{\partial \delta} > 0$ suggest increases in the discount factors will lead to increased savings and therefore increased capital formation ; $\frac{\partial k_0}{\partial \pi} > 0$ imply that as the proportion of rich increases, capital formation is higher; $\frac{\partial k_0}{\partial \psi} > 0$

suggests that an increase in the productivity of the poor leads to higher capital formation; $\frac{\partial k_1}{\partial \tau_1} < 0$ for $\beta > \theta$; $\frac{\partial k_2}{\partial \tau_2} < 0$ reflect the distortionary costs of redistribution.

Proposition 4 *(a) Capital is higher under FF than PAYG*

(b) For the same discount factor, $k_0 = k_1 = k_2$ *iff* $\tau_2 = 0$.

Proof. The Proof is in Appendix 3.1. ■

These results are consistent with the critiques of a PAYG system who suggest its existence reduces the capital stock (Feldstein, 1974, 1985).

3.4 The Planner's Problem

Given the competitive equilibrium, the planner's problem is to choose the tax rate τ so as to maximise $\{c_t^y, c_t^o, k_{t+1}\}_{t=0}^{\infty}$ subject to the allocation $\{y_t\}_{t=0}^{\infty}$. The planner weighs the utility of all agents living in the economy and each generation, born or unborn, is given a weight, $\rho \in (0, 1)$, which is decreasing in time. In the Benthamite tradition, we assume that the planner attaches the same weight to all agents living at a given point in time. Assuming the utility function remains the same across generations, the planner's social welfare function is thus:

$$W = \sum_{t=0}^{\infty} \rho^t \left(U\left(c^h\right)\right) \tag{35}$$

To enable a clearer comparison between the two modes of transfers, the analysis is carried out in steady state. This assumption enables us to avoid the issue of having to include the first generation of retirees (beneficiaries) and workers (taxpayers) when the intergenerational scheme is set up. Following Feldstein (1985), unless the future is discounted at a very high rate, the effect of the initial period over the long term will be relatively unimportant. Moreover, the steady state can reasonably be considered as the framework within which most of the agents will operate in an infinitely lived economy (Galor, 1992).

Normalising the first period's population to 1, the aim of the planner at t can be considered as choosing a sequence of $\{\tau, b/p\}_{t=0}^{\infty}$ so as to maximise welfare for all $t > 0$. The welfare function at t can thus be expressed as that of maximising the welfare of all living generations, young and old, rich and poor, simultaneously. Hence, $V_t = C = C^{y,h} + C^{o,h}$. This can be summarised in intensive form as:

$$V = \pi\left[\ln(c^{y,r}) + \beta \ln(c^{o,r})\right] + (1-\pi)\left[\ln(c^{y,p}) + \theta \ln(c^{o,p})\right] \quad (36)$$

The first part of the welfare function refers to the rich and their consumption over their lifetime whilst the second part refers to the consumption of the poor. Based on the optimal levels of consumption from the competitive equilibrium and the equilibrium conditions for the transfers from the planner, the welfare functions for the two modes of transfers can now be elaborated. For comparative purposes, we also include the welfare function without redistribution. We refer to the welfare functions as V_0, V_1 and V_2 and these are shown in Appendix 3.1.

V_0 represents welfare without redistribution whilst V_1 and V_2 represent welfare with redistribution in periods 1 (intragenerational) and 2 (intergenerational) respectively. If redistribution increases welfare such that the $V_1 \geq V_0$ and $V_2 \geq V_0$, the decision criteria for the planner can be summarised as follows:

(i) $V_1 - V_2 > 0$

If $V_1 - V_2 > 0$, the transfer in the first period yields the highest welfare for the economy and τ_1 is optimal. The system then approximates to a FF system, where each individual is responsible for providing for his own income in retirement.

(ii) $V_1 = V_2$

If $V_1 - V_2$, then both modes of transfer generate the same welfare and whether the redistribution takes place through τ_1 or τ_2 does not matter.

(iii) $V_2 - V_1 > 0$

If $V_2 - V_1 > 0$ then τ_2 is the optimal instrument to redistribute with. The transfer promotes a kind of a hybrid system where the rich fund their pensions under the FF scheme whilst the redistribution provides the poor with an additional source of income in retirement.

Before proceeding to calculate the optimal level of tax and the implications for the mode of redistribution, the impact of three of the key exogenous variables (π, ψ, n) for welfare are considered. For ease of manipulation, we consider V_0 and further assume that the agents have the same discount factor δ. The welfare function can then be summed up as:

$$V_0 = \left[\frac{(1+2\delta\alpha-\delta)}{(1-\alpha)}\right]\ln\left[\pi+(1-\pi)\psi\right]-\ln(1+n)+(1-\pi)\left(1+\delta\right)\ln\psi+z\left(\delta,\alpha\right)^{7} \tag{37}$$

Proposition 5 *(a) An increase in π increases welfare unambiguously.*

(b) An increase in ψ increases welfare unambiguously.

(c) An increase in n reduces welfare unambiguously

Proof. The Proof is in Appendix 3.1. ■

The results suggest that welfare is higher in richer economies or economies with relatively small differences in the productivity of the rich and poor. An increase in the population growth rate reduces capital per worker and as a result leads to a reduction in welfare.

3.4.1 The Optimal Level of Redistribution

The optimal level of redistribution or transfer τ from the rich to the poor can be found by setting $\frac{dV}{d\tau}=0$. Based on the resulting tax rates, we can then infer which type of redistribution is optimal and what type of mechanism will be put in place.

[7] $z\left(\delta,\alpha\right)=\delta\ln\alpha+\left[\frac{(1+\delta)\alpha}{(1-\alpha)}\right]\ln\delta-(2+\delta)\ln\left(1+\delta\right)+\left[\frac{1+\delta\alpha}{(1-\alpha)}\right]\ln(1-\alpha)$

The Golden Rule: Intragenerational Transfer The Golden Rule, where the returns to capital, $R = (1 + n)$, Samuelson's biological rate of interest, is considered. Steady state capital is given by eqn.(30). We replace $R_1 = (1 + n)$ in the welfare function V_1 and proceed to differentiate it with respect to τ_1. This results in:

$$\frac{\partial V_1}{\partial \tau_1} = \left[\begin{array}{c} \frac{-(1+\beta)}{(1-\tau_1)} + \frac{(1-\pi)(1+\theta)}{[(1-\pi)\psi+\pi\tau_1]} \\ + \left(\frac{[\pi(1+\beta)+(1-\pi)(1+\theta)]\alpha(\boldsymbol{\theta}-\boldsymbol{\beta})}{(1-\alpha)[\beta(1+\theta)\pi(1-\tau_1)+\theta(1+\beta)[(1-\pi)\psi+\pi\tau_1]]} \right) \end{array} \right]$$

We can see from the second part that there is a component of $\frac{\partial V_1}{\partial \tau_1}$ consisting of $(\theta - \beta)$. Hence we can consider $\frac{\partial V_1}{\partial \tau_1} = 0$ when all agents have the same discount factor. For $\beta = \theta = \delta$, $\frac{\partial V_1}{\partial \tau_1} = \left[\frac{-1}{(1-\tau_1)} + \frac{(1-\pi)}{([(1-\pi)\psi+\pi\tau_1])} \right] = 0$. Hence the optimal tax is:

$$\tau_1^* = (1 - \pi)(1 - \psi) \tag{38}$$

For $\beta = \theta$, the optimal level of transfer is simply a function of the relative gap in productivity between the rich and the poor and the proportion of poor in the economy.

Proposition 6 *An increase in π or ψ leads to a lower tax rate*

Proof. Follows from Eqn. (38) ■

This implies that as the proportion of rich in the economy increases, the optimal level of transfer from the rich to the poor falls - and this can be rationalised to the extent that, if there are fewer poor people in the economy, the contribution to be made by the rich for redistribution towards the poor falls relatively. Similarly, if the productivity gap between the rich and the poor falls, the optimal level of contribution by the rich declines. This can again be explained by the fact that as the wedge between the rich and the poor declines, the amount of funds needed to bridge the gap between the rich and the poor narrows.

$\tau_1^* \in (0,1)$ is always satisfied[8]. One of the clear implications is that the minimum rate of tax is achieved under conditions of near homogeneity. If all the agents are classed as rich ($\pi = \psi = 1$), then the optimal level of redistribution from young to rich is zero. Conversely, the lower the level of π and ψ, the higher the tax rate. This suggests that in a highly unequal economy with a small proportion of rich and where the poor have a lower productivity, the rich will have to bear a high burden for redistribution.

[8] For $\tau_1 > 0, \pi < 1$ and $\psi < 1$.For $\tau_1 < 1, \pi > 0$ and $\psi > 0$. Since the above conditions on π and ψ are always satisfied, it follows that $\tau_1^* \in (0,1)$.

General Case: Intragenerational Transfer We now consider the general case and proceed to find the optimal level of redistribution between the young rich and the poor. The steady state capital stock is defined as before by eqn.(30). The welfare function is shown in Appendix 3.1. The optimal tax is derived as previously and this yields:

$$\frac{\partial V_1}{\partial \tau_1} = \left[\begin{array}{c} \frac{-(1+\beta)}{(1-\tau_1)} + \frac{(1+\theta)(1-\pi)}{[(1-\pi)\psi+\pi\tau_1]} \\ + \left[\frac{[(1+2\beta\alpha-\beta)\pi+(1+2\theta\alpha-\theta)(1-\pi)](\boldsymbol{\theta}-\boldsymbol{\beta})}{(1-\alpha)[\beta(1+\theta)\pi+\theta(1+\beta)(1-\pi)\psi+[\theta-\beta]\pi\tau_1]} \right] \end{array} \right] \tag{39}$$

As for the Golden rule case, we can find that there is a $(\boldsymbol{\theta} - \boldsymbol{\beta})$ component present in eqn. (39) . We therefore consider the case for $\beta = \theta = \delta$ and this yields the same result as for the Golden rule:

$$\tau_1^* = (1-\pi)(1-\psi) \tag{40}$$

Proposition 7 *For intragenerational redistribution, the optimal tax rate is the same for the Golden rule and the general case.*

Proof. This can be seen by comparing eqns. (38) and (40) . ■

The Golden Rule: Intergenerational Transfer The case where the transfer is from young to old is now considered for the golden rule. The general welfare function and the first order condition are shown in Appendix 3.1. For the special case of $\beta = \theta = \delta$, $\frac{\partial V_2}{\partial \tau_2}$yields:

$$\frac{\partial V_2}{\partial \tau_2} = \left(\begin{array}{c} \frac{-1}{(1-\tau_2)} + \frac{(1-\pi)}{[(1-\pi)\psi+\pi\tau_2]} - \\ \left[\frac{\alpha}{(1-\alpha)[\pi(1-\tau_2)+(1-\pi)\psi]} + \frac{\alpha}{[(1+\delta)\alpha+\pi\tau_2(1-\alpha)]} \right] \end{array} \right) \tag{41}$$

Eqn. (41) does not allow for a significant degree of analytical tractability and we consider it again in the simulations section.

General Case: Intergenerational Transfer When redistribution takes place via the pensions scheme for the general case, capital is defined by eqn. (31). The welfare function is rather cumbersome and for $\beta = \theta = \delta$, differentiating V_2 with respect to τ_2 yields:

$$\frac{\partial V_2}{\partial \tau_2} = \left[\begin{array}{c} \left(\frac{-(1+\delta)}{(1-\tau_2)} \right) - \left(\frac{[1+2\delta\alpha-\delta]}{(1-\alpha)[(1-\pi)\psi+(1-\tau_2)\pi)]} \right) - \left(\frac{[(1+\delta)\alpha+((1+\delta)\pi-1)(\alpha-1)]}{((1+\delta)\alpha+\pi(1-\alpha)\tau_2)} \right) \\ + \left(\frac{(1+\delta)(1-\pi)[(1-\alpha)((1+\delta)(1-\pi)\psi+\delta\pi(1-2\tau_2))]}{[((1-\pi)\psi(1+\delta)\alpha)+\pi(1-\alpha)[(1+\delta)(1-\pi)\psi\tau_2+\delta\pi\tau_2(1-\tau_2))]]} \right) \end{array} \right] \tag{42}$$

Given the nature of eqn. (42), there is no straightforward closed form solution that is possible. Hence, we consider how τ_2 behaves as the parameter values change in the simulations section.

3.4.2 Impact of τ on utility of rich and poor

Given the nature of the utility functions, it is quite intuitive that the utility of rich (UR) is negative in taxes whilst that of the poor (UP) is positive. For ease of exposition, we will assume that the agents have the same discount factor. With redistribution in the first period, we can express the utility of the two types of agents as:

$$UR_1 = (1+\delta)\ln(1-\tau_1) + (1+\delta)\ln w_1 + \delta \ln R_1 \tag{43}$$

$$UP_1 = \ln\left(\left[\psi + \frac{\pi\tau_1}{(1-\pi)}\right] w_1\right) + \delta \ln\left(\left[\psi + \frac{\pi\tau_1}{(1-\pi)}\right] w_1 R_1\right) \tag{44}$$

Proposition 8 *The utility of the rich falls as τ rises whilst that of the poor increases.*

Proof. Since for $\beta = \theta = \delta$, k_1 is independent of τ_1, it follows that $\frac{\partial UR_1}{\partial \tau_1} < 0$ and $\frac{\partial UP_1}{\partial \tau_1} > 0$. ■

3.4.3 Simulations Results and Sensitivity Analysis

Having considered the behaviour of the tax rate under some specific settings, the optimal level and timing of redistribution is now considered in a general equilibrium framework encompassing capital, output, wages, interest rates, consumption, the tax rate and welfare. The optimal timing of redistribution will be based the decision criteria of the planner (as highlighted previously). For a given set of parameters and tax rates, the welfare function with the highest value determines in which of the two periods it is optimal to effect the transfer. Alternatively, there might be a set of parameters and tax rates for which it is not possible to improve on the initial distribution, in which case redistribution, in one form or another, is not optimal.

Choice of parameters The parameters are chosen to ensure the economy is dynamically efficient with an initial interest rate of 4 percent. The values for the population growth rates and the discount factors are in line with the range of values used for macro simulations (see de la Croix and Michel, 2002; Krueger and Kubler, 2006). In line with the demographic transitions affecting different regions of the world, there is a wide range of parameters that could have been used ranging from 0 for Europe to 2.4 percent for Africa with 1.3 percent for the world[9] (United Nations, 1999). For the purpose of the simulations we assume that population grows at a rate of 1 percent. The share of capital in production is set at 0.3. The discount factor of the rich

[9] Data pertain to 1995-2000.

(β) is 0.96, whilst that of the poor (θ) is initially 0.96 and it is then lowered to 0.9 for $\beta > \theta$. In the baseline, we assume that the poor earn 75 percent of that of the rich. The rich, π, make up 40 percent of the economy, with the poor accounting for the remaining 60 percent. ($n = 0.01; \alpha = 0.3, \beta = 0.96, \theta = 0.96/0.9, \psi = 0.75, \pi = 0.4$)

Simulation Results For the baseline, two cases for the discount factors are considered: Firstly, the discount factor of the poor (θ) and rich (β) is set to be the same at 0.96. In the second case, $\beta\,(0.96) > \theta\,(0.9)$ is considered. The baseline results are shown in Table 1.

The first case of $\beta = \theta$ is analysed for a general overview of the results. V_0, V_1 and V_2 are defined as previously. When redistribution takes place in the first period, for $\beta = \theta$, it can be seen that the capital stock is not (adversely) affected relative to the initial distribution and as a result, the output, wage and interest rate are the same. With intragenerational redistribution, the tax rate is 15% - and this confirms the earlier result that for $\beta = \theta$, the optimal level of transfer is defined by $\tau_1^* = (1-\pi)(1-\psi)$. There is an increase in aggregate welfare relative to the initial distribution. In this setting this takes

	α	β	θ	π	ψ	n		
$\beta = \theta$	0.3	**0.96**	**0.96**	0.4	0.75	0.01		
	k	**y**	**w**	**R**	UR	UP	$\boldsymbol{\tau}$	**Welfare**
V_0	0.1694	0.5870	0.4109	1.0397	-3.0639	-3.6278	0	-3.4022
V_1	0.1694	0.5870	0.4109	1.0397	-3.3824	-3.3824	0.1500	-3.3824
V_2	0.1468	0.5623	0.3936	1.1495	-3.2713	-3.4599	0.1059	-3.3845
	α	β	θ	π	ψ	n		
$\beta > \theta$	0.3	**0.96**	**0.9**	0.4	0.75	0.01		
	k	**y**	**w**	**R**	UR	UP	$\boldsymbol{\tau}$	**Welfare**
V_0	0.1652	0.5826	0.4078	1.0581	-3.0618	-3.5142	0	-3.3332
V_1	0.1647	0.5821	0.4075	1.0602	-3.3256	-3.3123	0.1260	-3.3176
V_2	0.1464	0.5619	0.3934	1.1513	-3.2301	-3.3822	0.0870	-3.3214

Table 1: Baseline Simulations

place through the following mechanism: (i) the utility of the rich falls; (ii) the utility of the poor rises and; (iii) all the agents consume the same amount after redistribution such that their utility is equalised.

When intergenerational redistribution takes place, the adverse impact of the redistribution on the capital stock results in a lower output and wage and higher interest rate. However, despite the distortion introduced in terms of capital formation, intergenerational redistribution still represents an improvement over the initial distribution. The same mechanism operates as regard consumption in that the rich consume less than the initial level whilst the poor consume more. However, since the tax rate is now smaller, the utility of the rich falls by a smaller amount whilst that of the poor rises by a smaller amount as well.

Whilst both V_1 and V_2 represent an improvement over V_0, it can be seen the welfare is higher for V_1 compared to V_2, suggesting that for the given set of parameters, intergenerational transfer is optimal. Each young agent would then fund for his own retirement and there is no need for pensions as a redistributive instrument in this set-up.

For $\beta > \theta$, the capital stock is now lower, compared to $\beta = \theta$, resulting in lower output and wages whilst the interest rate goes up. Aggregate welfare is higher than that of the initial distribution. There is a decrease in the tax rates compared to the case of $\beta = \theta$. As for $\beta = \theta$ welfare is unambiguously higher with redistribution, with welfare higher for intragenerational redistribution compared to intergenerational redistribution. For $\beta > \theta$ the optimal redistribution is thus still from young to young. One of the interesting aspects of the results is the fact that although $\tau_1$is almost $1\frac{1}{2}$ times τ_2, the welfare is still higher under intragenerational transfer implying that the welfare costs in terms of capital formation are fairly high with intergenerational transfer. The main results for the two cases for the given set of parameters are summarised in Table 2.

The results in Table 2 are in line with the previous findings that redistribution leads to a lower capital stock when the discount factor of the rich, β is greater than θ, the discount factor of the poor. As a result, $k^\delta > k$ which in turn yields $y^\delta > y; w^\delta > w$ and $R^\delta < R$. Though not shown in Table 2, with redistribution the consumption of the rich is always lower than without redistribution, whilst that of the poor is always higher.

Having considered the results we now proceed to show the impact of a 1% change in the exogenous parameters on the endogenous variables compared

(i)	(ii)	$(i)\ vs\ (ii)$
$\beta = \theta$	$\beta > \theta$	
$k_0^\delta = k_1^\delta > k_2^\delta$	$k_0 > k_1 > k_2$	$k^\delta > k$
$y_0^\delta = y_1^\delta > y_2^\delta$	$y_0 > y_1 > y_2$	$y^\delta > y$
$w_0^\delta = w_1^\delta > w_2^\delta$	$w_0 > w_1 > w_2$	$w^\delta > w$
$R_0^\delta = R_1^\delta < R_2^\delta$	$R_0 < R_1 < R_2$	$R^\delta < R$
$\tau_1^\delta > \tau_2^\delta$	$\tau_1 > \tau_2$	$\tau^\delta < \tau$
$V_0^\delta < V_2^\delta < V_1^\delta$	$V_0 < V_2 < V_1$	$V_0^\delta < V_0 < \mathbf{V}_2^\delta < \mathbf{V}_2 < V_1^\delta < V_1$

Table 2: Summary of Results for Selected Parameters

to the baseline. The results for $\beta > \theta$ are summarised in Table 3.

Changes in the exogenous parameters for $\beta = \theta$ and $\beta > \theta$ are unambiguous in so far as the impact on capital and consequently output, wages and interest rate are concerned. However, the main difference lies at the level of taxes and welfare. For $\beta = \theta$, the tax rate changes for young to young only when the proportions of rich and the level of productivity of the poor change since $\tau = (1 - \pi)(1 - \psi)$. Changes in the other parameters do not affect the optimal level of tax from young to young.

	k	**y**	**w**	**R**	UR	UP	**Tax**
α	$-$	$-$	$-$	$+$	$-$	$-$	$-$
β	$+$	$+$	$+$	$-$	$-$	$-$	$-$
θ	$+$	$+$	$+$	$-$	$-$	$-$	$-$
π	$+$	$+$	$+$	$-$	$+$	$+$	$-$
ψ	$+$	$+$	$+$	$-$	$+$	$+$	$-$

Table 3: Impact of a 1 percent change

Changes in α and n have a negative impact on capital. Changes in α has an unambiguous negative impact on welfare for both types of transfer. However, it has to be noted that the impact of n is infinitesimal compared to a change in α.

For the baseline case for $\beta > \theta$, capital is unambiguously and positively related to any change in β, θ, π and ψ. A 1% increase in β or θ unambiguously reduces welfare. This could possibly be due to the fact that at high levels of β and θ, any increase in the parameters causes a higher decrease in the utility from the reduced consumption than the gain from the increase in capital.

The relationships between capital (hence, output and wages) and both π and ψ are positive. An increase in π has a positive impact on welfare and is consistent with the view that a homogeneous economy, in terms of having more rich people, requires a lower level of transfer from the rich to the poor. In the same vein, an increase in the proportion of rich people in the economy leads to an increase in welfare. The same reasoning applies to ψ - an increase in productivity leads to a reduction in the gap between the rich and the poor and as a result a lower level of transfer is required whilst overall welfare is higher.

Sensitivity Analysis

Changes in the discount factor: So far, consistent with Pigou's "faulty telescopic faculty", it has been assumed that agents suffer from partial myopia in that they discount the future. Starting with the case where agents do not discount the future, that is they have perfect foresight, a combination of cases where agents discount the future at high rates are considered. Except for β and θ, all the other parameters are taken to be the same as in the previous experiments.

$\boldsymbol{\beta} = \boldsymbol{\theta} = \mathbf{1}$: When none of the agents discount the future, this can be

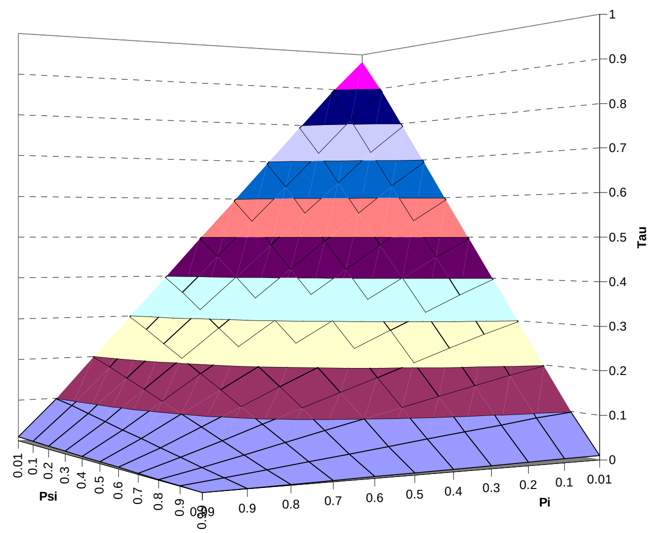


Figure 5: Tau for given pi and psi

summarised as $\beta = \theta = 1$. When the transfer is from young to young, the general result $\tau_1 = (1 - \pi)(1 - \psi)$ still applies whilst if there is an intergenerational transfer, the optimal level of tax is lower. This also results in a higher level of capital. Compared to the initial baseline, aggregate welfare is *lower* in all three cases. Figure 5 shows all the possible intragenerational tax rates for $\tau_1 = (1 - \pi)(1 - \psi)$.

$\boldsymbol{\beta = 1; \theta \to 0}$: Assuming that the rich do not discount the future, we initially begin with a similar discount factor for the poor. We then proceed

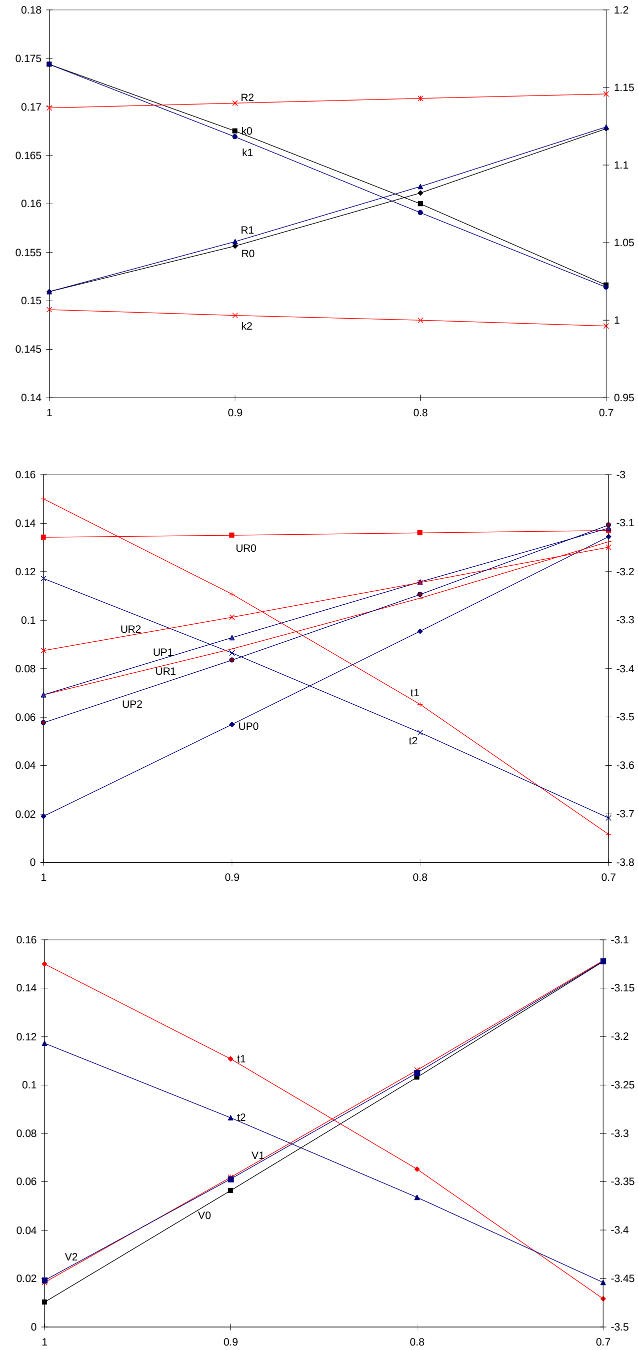


Figure 6: Impact of Discount Factor on Welfare (Right), Capital and Tax Rates (Left)

by allowing the poor to discount the future at a higher rate. We find that as the wedge increases marginally, this leads to an increase in the level of welfare under all settings. The fact that welfare increases initially as the poor discount the future at a higher level can be attributed to the fact that the initial increase in consumption boosts welfare by a higher level than the fall in capital. However, as the wedge widens, the associated costs in terms of lower capital and higher taxes dominate. The impact of the wedge on welfare is summarised in Figure 6.

It can be noticed that when both β and θ are close to 1, the optimal redistribution scheme is intergenerational, albeit V_2 is marginally greater than V_1. This result is not consistent with the Aaron condition in that when $\beta = \theta = 1$, the interest rate for $R_1 = 1.0185$ and $R_2 = 1.1369$; implying that both economies are dynamically efficient ($n = 0.01$).

However, once $\theta < 0.9$, intragenerational transfer is optimal throughout. As θ falls further we note that the tax rates under both schemes fall. For $\theta < 0.6$, it is not optimal to have any redistribution scheme for this set of parameters. This is due to the fact that a fall in the discount rate causes a fall in capital. The decrease in the tax rate mitigates the impact of the fall but does not fully compensate for it. For $\theta < 0.6$, the cost in terms of capital is so high that the optimal tax rate is zero.

Changes in proportions of rich and productivity of poor: We have seen that the proportions of rich and the productivity of the poor will play an important role in determining welfare in this set-up. An economy with a low proportion of rich has a lower level of utility (relative to baseline) and similiarly, as the gap between the rich and the poor widens, this has detrimental effects on welfare. Any changes in π or ψ will affect capital and hence welfare. Keeping all the other parameters unchanged relative to the baseline, π and ψ are allowed to vary and the resulting impact on the policy variables considered. The general equilibrium results for $\beta > \theta$ are summarised in Figure 7.

The first set of diagrams represent the evolution of π, the proportion of rich in the economy. We find that when π is set to 1% it leads to lower capital and hence the interest rate is higher, nearing 20% compared to 4% for the baseline (output and wages are also lower). The wedge between capital under various modes of redistribution persists such that $k_0 > k_1 > k_2$. We also find that for low levels of π, the tax rate is higher, suggesting that the

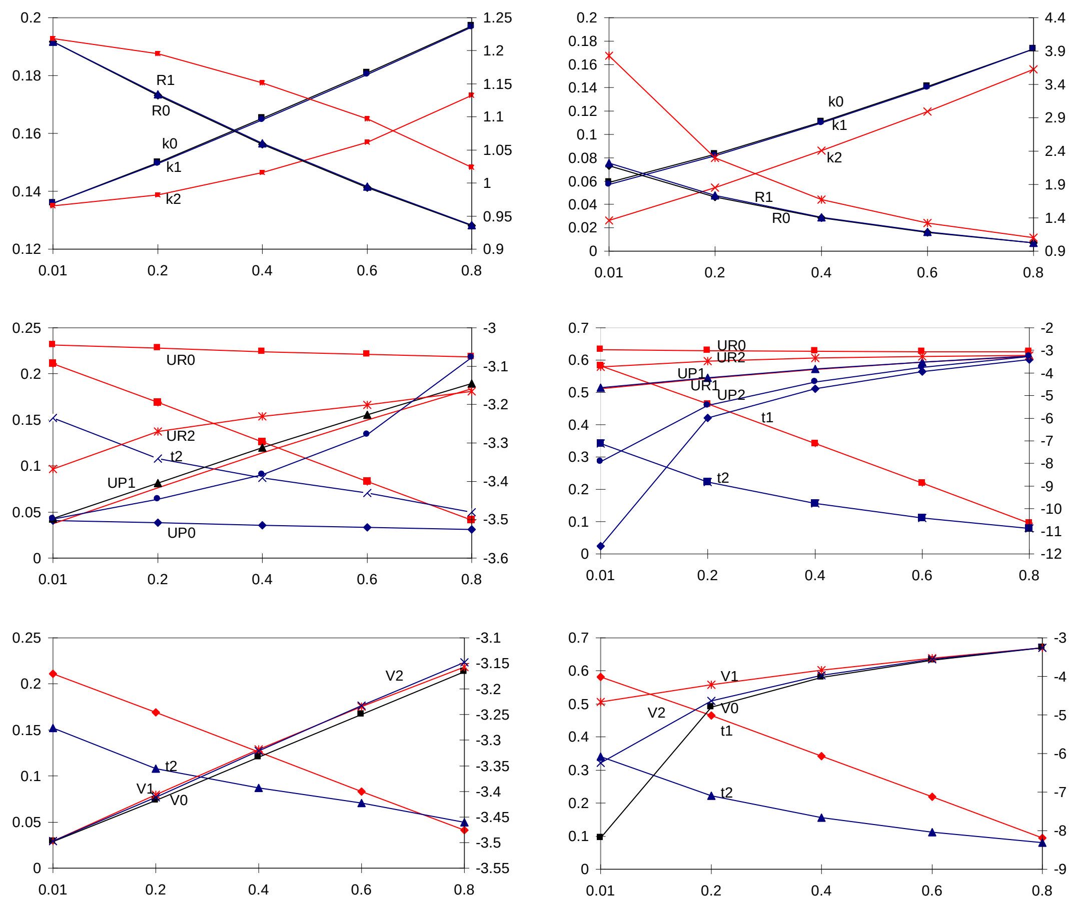


Figure 7: Impact of Pi (R) and Psi (L)

higher the inequality, the higher the level of redistribution undertaken. As π increases, we find that capital increases and the interest rates falls such that dynamic inefficiency ($r < n$) creeps in beyond a certain level. Dynamic inefficiency creeps in faster in the case of intragenerational transfers (since $k_1 > k_2$). Consistent with our finding that more equal societies require less redistribution, the tax rate falls gradually. Welfare increases unambiguously.

The same mechanism applies to the relative productivity of the poor, ψ. When the productivity of the poor is low it severely reduces capital. Welfare is the lowest among all the experiments considered and the tax rate is highest as well. We however find that as ψ starts to increase, capital increases consistently, with the wedge between k_2 still apparent, and the interest rates start to converge and fall (with dynamic inefficiency beyond a certain level). There is a convergence and increase in welfare as ψ increases with the convergence faster at lower levels of ψ.

Once again we find that the Aaron condition is violated in that there is a range of parameters for which it is optimal to have intergenerational redistribution in a dynamically efficient economy. It is noted that when π and ψ exceed a certain level, which is relatively high, it is no longer optimal to have redistribution from the rich to the poor.

Changes in population growth rates: As far as n is concerned, its impact in the current set up is relatively subdued in the sense that there have to be significant changes in the population growth rate for there to be any significant impact on the variables under consideration. We consider changes ranging from a 1% fall to a 1.5% increase in population from one period to the next. A fall in n is associated with an ageing population. The findings are in line with expectations in that as the population growth rate increases, capital available per worker falls and the interest rate rises. The wedge in capital is still apparent. The impact on the tax rate is infinitesimal though it does tend to fall as the population growth rate increases. In the case of the population growth rate increasing from -1.5% to 1%, for intragenerational transfer, the fall is only 1 basis point whilst in the case of intergenerational transfer, it is 13 basis points. Welfare falls as the population growth rate increases due to the decline in capital per worker.

3.4.4 When is pensions an optimal tool for redistribution?

A transfer from the rich to the poor through intragenerational taxes (young-to-young) is almost always welfare improving. The same applies to transfers through pensions with some restrictions in the sense that there are certain instances where it is optimal not to have pensions.

We now consider under what conditions pensions can be an optimal instrument for redistribution. For ease of exposition, the parameter values are set as before and we only consider how the decision is affected by changes in the proportions of rich and the level of productivity of the poor; the two parameters that seem to have the highest incidence on the optimal timing of redistribution. For pensions to be the optimal redistribution instrument, we require, $V_2 - V_1 > 0$.

Figure 8 depicts how the optimal instrument varies as the proportions of rich and productivity of the poor change. In the baseline π had been set at 0.4 and for the given set of parameters, intragenerational transfer was optimal. Allowing π to increase results in a bridging of the gap between V_1 and V_2. Intragenerational transfer is optimal for all $\pi < 0.54$. For $\pi = 0.54$, $V_1 = V_2$ and two modes of transfers yield the same welfare. If $\pi > 0.54$, then an intergenerational transfer is optimal. It can also be seen that for higher levels of ψ intergenerational redistribution is preferred.

The same mechanism applies to ψ which had been set at 0.75 in the baseline and for the given set of parameters, intragenerational transfer was opti-

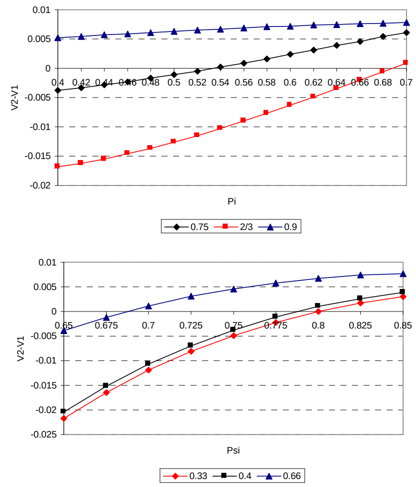


Figure 8: Optimal Timing of Redistribution

mal. By allowing ψ to increase, intragenerational transfer remains optimal for $\psi < 0.79$. increases to around 79% when welfare is the same irrespective of the mode of redistribution in place. For $\psi = 0.79$, $V_1 = V_2$ and two modes of transfers yield the same welfare. For $\psi > 0.79$, intergenerational transfer is optimal. Higher levels of π lead to intergenerational transfer being preferred.

In so far, we have assumed that either π or ψ vary, but that they do not vary simultaneously. For instance, if π falls to 0.01, ψ needs to increase to around 0.815 to ensure that $V_1 = V_2$. If ψ is higher (lower), then it is optimal to transfer from young-to-old (young-to-young). On the other hand if π increases to 90%, ψ can be as low as 14.7% and yet $V_1 = V_2$. Conversely, if ψ is 1%, π needs to be as high as 91.55% to ensure $V_1 = V_2$. If we allow ψ

to be 90%, then intergenerational redistribution is optimal for any π.

The mode of transfer will also determine the type of pensions scheme in place. With intergenerational transfer, there is a hybrid system with the rich saving through a FF scheme through the provision of their own pension whilst the poor, although they might save also receive an additional transfer when old in the form of a PAYG scheme. On the other hand, with intragenerational transfer, all redistribution takes place in the first period and every individual is responsible for the provision of his or her own pension in retirement. The system is FF in that case.

The above results suggest that it is optimal to redistribute in the first period if an economy is characterised by pervasive inequality and the productivity of the poor relative to the rich is low. Under those circumstances, redistributing in the first period yields higher welfare than redistributing through pensions. However, in an economy where inequality is relatively low, transferring through the pensions scheme is optimal. We have seen that for most part the results are in line with the Aaron condition, in that when $r < n$, intergenerational redistribution is preferred. However, there is a range in a dynamically efficient economy when it is optimal to redistribute through pensions. This ensures the economy remains dynamically efficient. Without the intergenerational redistribution, the economy would tend towards dynamic inefficiency.

3.5 An Application

We now estimate some of the parameters for 39 countries which is made up as follows: Europe (14) , America (8), Africa (10), Asia (5) and Pacific (2). The aim is to find out some estimates for the parameters α, β, π, ψ, and n and then consider which type of redistributive scheme would be in place. For simplicity, we assume that the discount factors of the poor and the rich are the same such that $\beta = \theta$. The data definitions and sources are provided in Appendix 3.2.

The initial results can be summarised as follows: Redistribution unambiguously increases welfare in all the countries sampled. Out of the 39 countries, intragenerational tax is optimal in 11 whilst intergenerational is optimal in 28. There are two important findings: There is always a positive level of intragenerational tax that exists, however for some values of α, the optimal intergenerational tax is zero. In all cases where intragenerational taxes were chosen, the economies were dynamically efficient. Of the 28 economies where

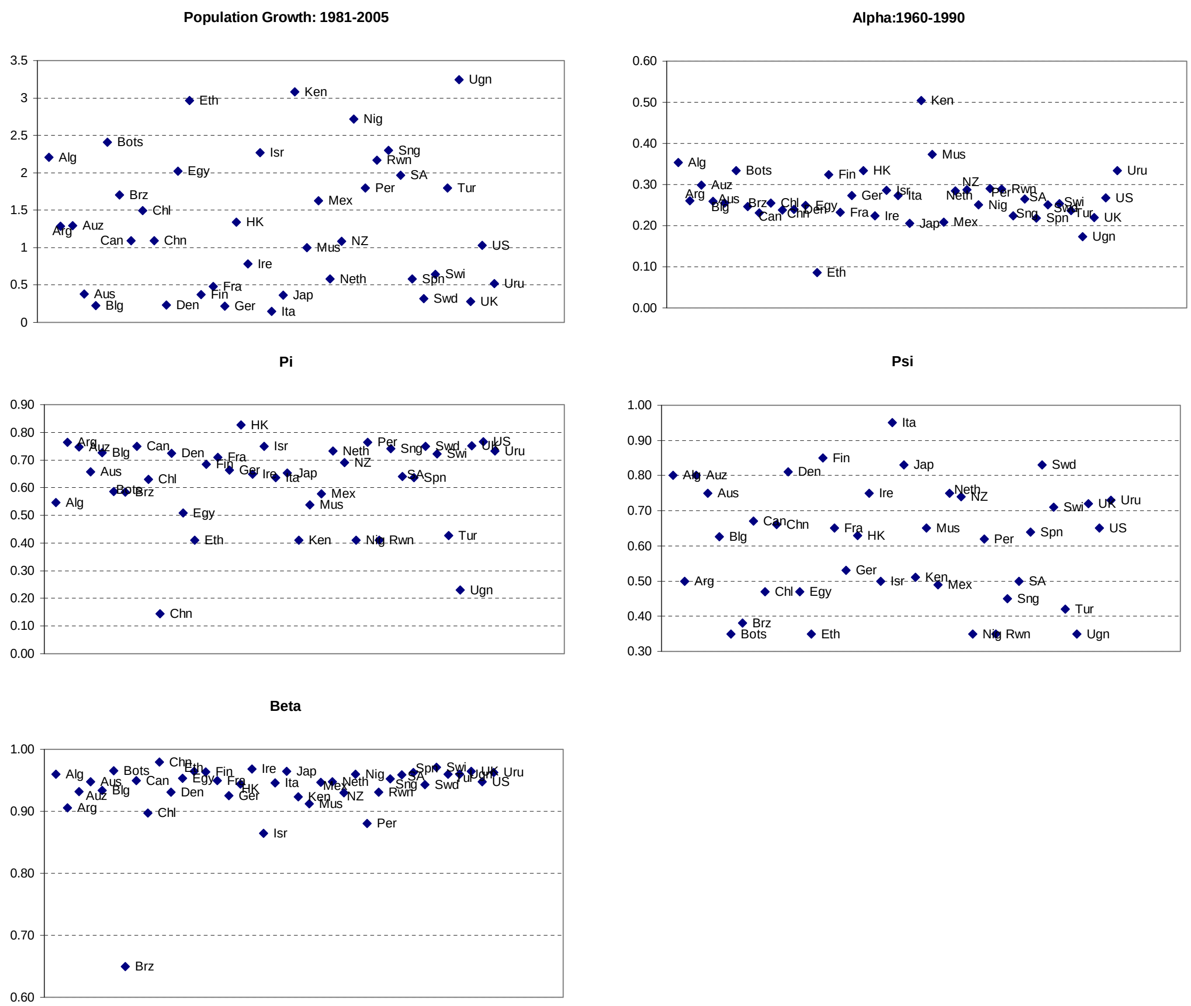


Figure 9: Overview of Estimated Exogenous Parameters

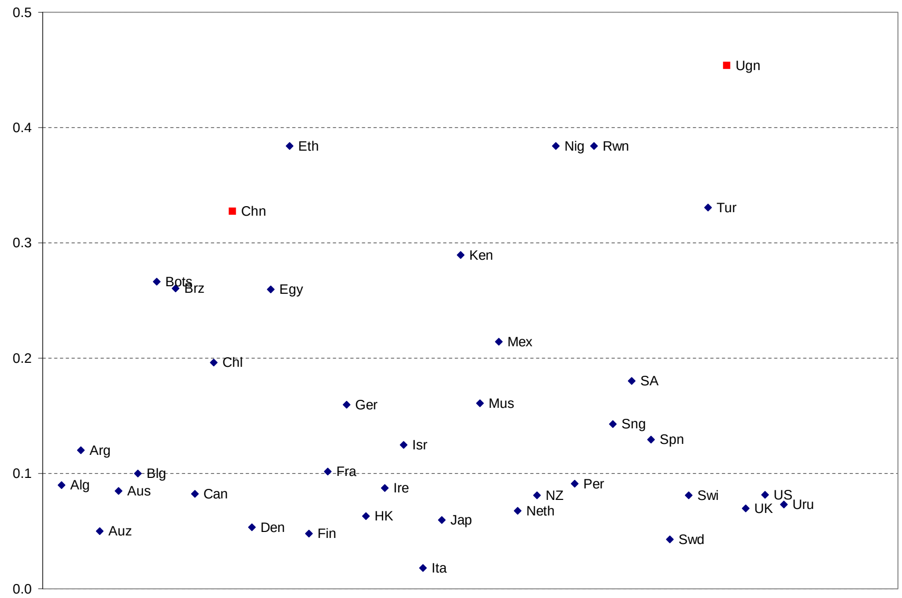


Figure 10: Overview of Optimal Tax Rates

intergenerational taxes was chosen, 11 became dynamically efficient after redistribution. However, for 17 countries the dynamic inefficiency persisted. After adjusting the α parameter upward above a certain level to ensure dynamic efficiency, it was found that in 37 of the economies intragenerational redistribution is optimal. The optimal tax rates in a dynamically efficient context are shown in Figure 10.

When we consider the adjusted parameters of α for which all the economies are dynamically efficient, we note that intragenerational redistribution takes place in 37 cases with an average tax of 14.6% whilst intergenerational tax is restricted to China and Uganda with tax rates of 32.7% and 45.5% respectively. We also note that the tax rate is highest in developing countries which is in line with the hypothesis that these countries have a low proportion of workers in services and the disparity between the rich and the poor in terms of wages (productivity) is highest. We note that for most of the developing countries, the tax rate is in the range of 20-50% whilst for the developed nations, it is under 20%.

3.6 Conclusions and Extensions

We have considered the impact of redistribution in an overlapping generations economy and considered some of the implications in a general equilibrium framework. We adopt a relatively underutilised approach to consider the pensions scheme as a redistribution instrument as opposed to the more conventional income support. Our results confirm some of the earlier findings on redistribution that it can be costly and there are distortions that arise in the economic decision making of the recipients. In our framework, this takes the form of a crowding out of capital resulting in an adverse impact on output and wages whilst causing interest rate to rise. We also find that the level and timing of redistribution matters as well. Whilst the poor benefit unambiguously from redistribution, the rich are worse off. Our other findings suggest that in cases of high inequality, that is, when there is a very small proportion of rich and the poor have a low productivity, such an economy is most likely to be dynamically efficient and intragenerational transfer is the optimal instrument. On the other hand, when inequality is fairly low, then the economy can tend towards dynamic inefficiency and it is then optimal to redistribute though intergenerational transfers. Though redistribution increases welfare in most cases, in some situations, it is optimal to have no redistribution. We have seen that for most part the results are in line with the Aaron condition, in that when $r < n$, intergenerational redistribution is preferred. However, there is a range in a dynamically efficient economy when it is optimal to redistribute through pensions.

Using a relatively simple framework, we have shown how the optimal method of redistribution might vary as the productivity and the relative proportions of the agents vary. However, we might enrich the environment by including two additional factors. We have so far assumed that there is only two types of agents, rich or poor. However, it might be appropriate to include a continuum of agents that differ according to their productivity and see how this affects the timing and level of redistribution. Moreover, in so far, we have considered the case that redistribution is carried out only through income support or pensions, that is, it is restricted to one period. This could be extended to include a situation whereby a proportion of the redistribution takes place in the first period whilst the remainder takes place in the second period. This is considered in the next chapter.

Appendix 3.1

Welfare functions for V_0, V_1 and V_2.

$$
\begin{aligned}
V_0 \;&=\; \pi\left[(1+\beta)\ln\left(\frac{w_0}{1+\beta}\right)+\beta\ln\left(\beta R_0\right)\right] \\
&\quad +(1-\pi)\left[(1+\theta)\ln\left(\frac{\psi w_0}{1+\theta}\right)+\theta\ln\left(\theta R_0\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
V_1 \;&=\; \pi\left[(1+\beta)\ln\left(\frac{(1-\tau_1)w_1}{1+\beta}\right)+\beta\ln\left(\beta R_1\right)\right] \\
&\quad +(1-\pi)\left[(1+\theta)\ln\left(\frac{w_1}{1+\theta}\left[\psi+\frac{\pi\tau_1}{(1-\pi)}\right]\right)+\theta\ln\left(\theta R_1\right)\right]
\end{aligned}
$$

$$
\begin{aligned}
V_2 \;&=\; \pi\left[(1+\beta)\ln\left(\frac{(1-\tau_2)w_2}{1+\beta}\right)+\beta\ln\left(\beta R_2\right)\right] \\
&\quad +(1-\pi)\left[(1+\theta)\ln\left(\frac{w_2}{1+\theta}\left[\psi+\frac{\pi(1+n)\tau_2}{(1-\pi)R_2}\right]\right)+\theta\ln\left(\theta R_2\right)\right]
\end{aligned}
$$

Golden Rule: Intragenerational transfers Replacing $R_1 = (1+n)$ in the welfare function V_1 now results in:

$$V_1 = \left[\begin{array}{c} \pi\left(1+\beta\right)\ln(1-\tau_1) + (1-\pi)\left(1+\theta\right)\ln\left[(1-\pi)\psi+\pi\tau_1\right] + z_1\left(\beta,\theta,\alpha,\pi,n\right) + \\ \frac{[\pi(1+\beta)+(1-\pi)(1+\theta)]\alpha}{(1-\alpha)}\ln\left[\beta\left(1+\theta\right)\pi(1-\tau_1)+\theta\left(1+\beta\right)\left[(1-\pi)\,\psi+\pi\tau_1\right]\right] \end{array} \right]$$

where z_1 is a set of parameters independent of the policy term.

Golden Rule: Intergenerational transfers

$$V_2 = \left[\begin{array}{c} \pi\left(1+\beta\right)\ln(1-\tau_2) + (1-\pi)\left(1+\theta\right)\ln\left[(1-\pi)\psi+\pi\tau_2\right] \\ +\frac{[\pi(1+\beta)+(1-\pi)(1+\theta)]\alpha}{(1-\alpha)}\ln\left(\frac{[\beta(1+\theta)\pi(1-\tau_2)+\theta(1+\beta)(1-\pi)\psi]}{[(1+\theta)\alpha+\pi\tau_2(1-\alpha)]}\right) + z_2^*\left(\beta,\theta,\alpha,\pi\right) \end{array} \right]$$

where z_2^* represents a set of parameters independent of the policy term.

The optimal tax is now represented by:

$$\frac{\partial V_2}{\partial \tau_2} = \left[\begin{array}{c} \frac{-(1+\beta)}{(1-\tau_2)} + \frac{(1-\pi)(1+\theta)}{[(1-\pi)\psi+\pi\tau_2]} - \\ \left[\frac{\beta\alpha(1+\theta)[\pi(1+\beta)+(1-\pi)(1+\theta)]}{(1-\alpha)[\beta(1+\theta)\pi(1-\tau_2)+\theta(1+\beta)(1-\pi)\psi]} + \frac{[\pi(1+\beta)+(1-\pi)(1+\theta)]\alpha}{[(1+\theta)\alpha+\pi\tau_2(1-\alpha)]}\right] \end{array} \right]$$

General Case: Intragenerational Transfer The welfare function V_1 can be summed up as:

$$V_1 = \left[\begin{array}{c} \left(1+\beta\right)\pi\ln(1-\tau_1) + \left(1+\theta\right)\left(1-\pi\right)\ln\left[(1-\pi)\psi+\pi\tau_1\right] + \\ \left[\frac{(1+2\beta\alpha-\beta)\pi+(1+2\theta\alpha-\theta)(1-\pi)}{(1-\alpha)}\right]\ln\left(\beta\left(1+\theta\right)\pi+\theta\left(1+\beta\right)\left(1-\pi\right)\psi+\left[\theta-\beta\right]\pi\tau_1\right) + z_2 \end{array} \right]$$

where $z_{\mathbf{2}}$ is a set of parameters independent of the policy term.

General Case: Intergenerational Transfer The welfare function now takes the following form:

$$V_2 = \left[\begin{array}{c} (1+\beta)\,\pi \ln(1-\tau_2) + \{A\} \ln\left[D - \beta\,(1+\theta)\,\pi\tau_2)\right] - \{B\} \ln\left[(1+\theta)\,\alpha + \pi(1-\alpha)\tau_2\right] + \\ (1+\theta)\,(1-\pi) \ln\left[E + \pi(1-\alpha)\left[(1+\beta)\,(1-\pi)\psi + \beta\pi\right]\tau_2 - \pi(1-\alpha)\beta\pi\tau_2^2)\right] \end{array} \right]$$

where $A = \frac{1}{(1-\alpha)} \left\{ \left[\begin{array}{c} \left[(1+\beta)\,\pi + (1+\theta)\,(1-\pi)\right]\alpha + \left[\beta\pi - (1-\pi)\right](\alpha-1) \\ -\,(1+\theta)\,(1-\pi)\,(1-\alpha) \end{array} \right] \right\}$;

$B = \frac{1}{(1-\alpha)} \left\{\left[(1+\beta)\,\pi + (1+\theta)\,(1-\pi)\right]\alpha + \left[\beta\pi - (1-\pi)\right](\alpha-1)\right\}$;

$D = \left[\theta\,(1+\beta)\,(1-\pi)\psi + \beta\,(1+\theta)\,\pi\right]$;

$E = (1-\pi)\psi\,(1+\beta)\,\alpha$,

are sets of parameters independent of the policy term.

The optimal level of tax now results in:

$$\frac{\partial V_2}{\partial \tau_2} = \left[\begin{array}{c} \frac{-(1+\beta)}{(1-\tau_2)} - \{A\} \frac{\beta(1+\theta)}{[D-\beta(1+\theta)\pi\tau_2)]} - \frac{[(1+\beta)\pi+(1+\theta)(1-\pi)]\alpha+[\beta\pi-(1-\pi)](\alpha-1)}{[(1+\theta)\alpha+\pi(1-\alpha)\tau_2]} \\ + \frac{(1+\theta)(1-\pi)(1-\alpha)[(1+\beta)(1-\pi)\psi+\beta\pi(1-2\tau_2)]}{\left[(1-\pi)\psi(1+\beta)\alpha+\pi(1-\alpha)[(1+\beta)(1-\pi)\psi+\beta\pi]\tau_2-\pi(1-\alpha)\beta\pi\tau_2^2)\right]} \end{array} \right]$$

Impact of τ_2 on Utility of Rich and Poor. We carry out the same procedure for τ_2. With one major difference, in that τ_2 now influences capital even if the agents have the same discount factor. The utility of the rich and the poor is given by:

$$UR_2 = [(1+\delta)\ln(1-\tau_2) + (1+\delta)\ln w_2 + \delta \ln R_2]$$

$$UP_2 = (1+\delta)\ln\left([(1-\pi)R_2\psi + \pi(1+n)\tau_2] + \ln w_2 - \ln(1-\pi)R_2\right) + \delta \ln R_2$$

It then follows that:

$$\frac{\partial UR_2}{\partial \tau_2} = \left(\begin{array}{c} \frac{-(1+\delta)}{(1-\tau_2)} \\ + \left[\frac{((1+2\delta)\alpha-\delta)\pi}{(1-\alpha)}\right] \left[\frac{(1-\alpha)}{[(1+\delta)\alpha+\pi\tau_2(1-\alpha)]} - \frac{1}{[\pi(1-\tau_2)+(1-\pi)\psi]}\right] \end{array} \right)$$

$$\frac{\partial UP_2}{\partial \tau_2} = \left(\begin{array}{c} \left[\frac{(1+\delta)(1-\alpha)[(1-\pi)\psi(1+\delta)+\delta\pi(1-2\tau_2)]}{[(1-\pi)\psi[(1+\delta)\alpha+\pi\tau_2(1-\alpha)]+\pi\delta(1-\alpha)[(1-\pi)\psi+\pi(1-\tau_2)]\tau_2]}\right] \\ - \left[\frac{((1+2\delta)\alpha-\delta)}{(1-\alpha)[\pi(1-\tau_2)+(1-\pi)\psi]} + \frac{(1+\delta\alpha)}{[(1+\delta)\alpha+\pi(1-\alpha)\tau_2]}\right] \end{array} \right)$$

The utility of the rich falls and that of the poor increases if $\frac{\partial UR_2}{\partial \tau_2} < 0$ and $\frac{\partial UP_2}{\partial \tau_2} > 0$.

Proof of Proposition 3:

$(1a)\ \frac{\partial k_0}{\partial \delta} = \frac{\partial k_1}{\partial \delta} = \frac{1}{1-\alpha}\left[\frac{(1-\alpha)\{\pi+(1-\pi)\psi\}}{(1+n)(1+\delta)^2}\right]^{\frac{\alpha}{1-\alpha}} > 0$ Eqns. $(32)-(33)$

$(1b)\ \frac{\partial k_0}{\partial \pi} = \frac{\partial k_1}{\partial \pi} = \frac{1}{1-\alpha}\left[\frac{\delta(1-\alpha)(1-\psi)}{(1+n)(1+\delta)}\right]^{\frac{\alpha}{1-\alpha}} > 0$ Eqns. $(32)-(33)$

$(1c)\ \frac{\partial k_0}{\partial \psi} = \frac{\partial k_1}{\partial \psi} = \frac{1}{1-\alpha}\left[\frac{\delta(1-\alpha)(1-\pi)}{(1+n)(1+\delta)}\right]^{\frac{\alpha}{1-\alpha}} > 0$ Eqns. $(32)-(33)$

$(1d)\ \frac{\partial k_0}{\partial n} = \frac{\partial k_1}{\partial n} = \frac{1}{1-\alpha}\left[\frac{-\delta(1-\alpha)[\pi+(1-\pi)\psi]}{(1+\delta)(1+n)^2}\right]^{\frac{\alpha}{1-\alpha}} < 0$ Eqns. $(32)-(33)$

$(1e)$ Comparing eqns. (30) and (33), we can see that eqn. (33) does not depend on τ_1. Moreover, eqn. $(32) =$ eqn. (33).

Else,$\frac{\partial k_1}{\partial \tau_1} = \frac{1}{1-\alpha}\left[\frac{(1+n)(1+\beta)(1+\theta)(1-\alpha)\pi[\boldsymbol{\theta}-\boldsymbol{\beta}]}{[(1+n)(1+\beta)(1+\theta)]^2}\right]^{\frac{\alpha}{1-\alpha}} < 0$ for $\beta > \theta$ eqn. (30)

$(1f)\ \frac{\partial k_2}{\partial \tau_2} = \frac{1}{1-\alpha}\left[-\frac{\delta(1+n)(1-\alpha)\alpha\pi\{(1+\delta)\alpha+(1-\alpha)[\pi+(1-\pi)\psi]\}}{[(1+n)[(1+\delta)\alpha+\pi\tau_2(1-\alpha)]]^2}\right]^{\frac{\alpha}{1-\alpha}} < 0$ eqn. (34)

Proof of Proposition 4:

$(2a)$ For δ, $k_1 - k_2$ $[Eqn.(33) - Eqn.(34)]$ simplifies to the following condition:

$$\delta\pi\tau_2\left[\pi(1-\alpha)+(1-\pi)(1-\alpha)\psi+\alpha\,(1+\delta)\right] > 0$$

It follows that for all $\tau_2 > 0$, the above condition is always satisfied and hence $k_1 > k_2$.

$(2b)$ If $\tau_2 = 0$, in eqn. (34), then eqn. $(32) - (34)$ are the same.

Proof of Proposition 5:

$$
\begin{aligned}
(3a) &\quad : \quad \frac{\partial V_0}{\partial \pi} = \left[\frac{(1+2\delta\alpha-\delta)(1-\psi)}{(1-\alpha)[\pi+(1-\pi)\psi]}\right] - (1+\delta)\ln\psi > 0 \\
(3b) &\quad : \quad \frac{\partial V_0}{\partial \psi} = \left[\frac{(1+2\delta\alpha-\delta)(1-\pi)}{(1-\alpha)[\pi+(1-\pi)\psi]}\right] + (1-\pi)(1+\delta)\frac{1}{\psi} > 0 \\
(3c) &\quad : \quad \frac{\partial V_0}{\partial n} = \frac{-1}{(1+n)} < 0
\end{aligned}
$$

Proof of Proposition 8:

$$
\begin{aligned}
\frac{\partial UR_1}{\partial \tau_1} &= \frac{-(1+\delta)}{(1-\tau_1)} < 0 \\
\frac{\partial UP_1}{\partial \tau_1} &= \frac{(1+\delta)\pi}{[(1-\pi)\psi + \pi\tau_1]} > 0
\end{aligned}
$$

Appendix 3.2

Europe	**America**	**Africa**	**Asia**
(14)	(8)	(10)	(5)
Austria	Canada (Can)	Algeria (Alg)	China (Chn)
Belgium (Blg)	United States (US)	Botswana (Bots)	Hong Kong (HK)
Denmark (Den)	Argentina (Arg)	Egypt (Egy)	Japan (Jap)
Finland (Fin)	Brazil (Brz)	Ethiopia (Eth)	Singapore (Sng)
France (Fra)	Chile (Chl)	Kenya (Ken)	Israel (Isr)
Germany (Ger)	Mexico (Mex)	Mauritius (Mus)	
Ireland (Ire)	Peru (Per)	Nigeria (Nig)	
Italy (Ita)	Uruguay (Uru)	Rwanda (Rwn)	**Pacific**
Netherlands (Neth)		South Africa (SA)	(2)
Spain (Spn)		Uganda (Ugn)	Australia (Auz)
Sweden (Swd)			New Zealand (NZ)
Switzerland (Swi)			
Turkey (Tur)			
United Kingdom (UK)			

$\boldsymbol{\alpha}$

α is the share of capital in output. We use the database of Nehru and Dhareshwar (1995) from the World Bank. The period covers 1960-1990. We define $\alpha = \frac{\text{Physical Capital Stock}}{GDP}$ and the initial estimates vary from 0.09 for Ethiopia to 0.5 for Kenya with an average of 0.27 for the sample.

$\boldsymbol{\beta}$

$\beta = \frac{1}{1+r}$ where r is the annual real rate of interest for the period 1990 to 2005. However, where estimates were not available, they were assigned a value of 0.96, in line with the literature. The estimates ranged from 0.65 for Brazil to 0.98 for China with an average of 0.94 for the sample. (Source: IFS)

π

π was derived from ILO data and considered to be the proportion of workers working in the services sector. The values here ranged from 0.14 for China to 0.83 for Hong Kong with a mean of 0.62.

ψ

ψ was derived from ILO data as follows:

1. The average wage in the economy throughout the sectors was derived.
2. The workers were then classed into 2 categories - low and high (below and above average wage);

3. The ratio of the wage of the low worker relative to the high worker was then taken to be ψ

In this case, the mean was 0.61 within a range of 0.35 for Uganda and 0.95 for Italy.

n

n is the annual average of the population growth rate from 1981-2005. (Source: WDI)

4 REDISTRIBUTION-TILTING WITH PLANNER INEQUALITY AVERSION

4.1 Introduction

Governments intervene in economies to achieve a multiplicity of objectives. Among others, carrying out redistribution from the rich to the poor has been one of the main reasons for intervention. However, redistribution often entails costs, be they budgetary or in terms of the disincentives it gives rise to. Nonetheless, governments are called to carry out redistribution in various forms. As such, an optimal redistribution scheme can help in defining what is the optimal level of redistribution and when it should take place. This chapter is an extension of chapter 3 whereby we considered whether redistribution should be intergenerational or intragenerational. The planner thus had only one policy tool to choose from, i.e., either income support where all the resources are transferred to the young or pensions whereby the poor benefit from all the redistribution.

The basic gist of the current paper remains the same: redistribution takes place from the young rich to the poor. However, unlike the previous chapter where redistribution could be either to the young or the old, this time round we allow for redistribution to be to both the young and the old. The basic question focuses on what determines the proportion which is allocated to the young and the old. If all the tax raised from the young rich is transferred to the young poor, then all agents fund for their own retirement and we have a fully-funded pensions system. On the other hand, if all the money is allocated to the old poor, we have a kind of a hybrid system whereby the rich fund for their own pensions whilst for the poor, the pay-as-you-go (PAYG) pensions is targeted and is purely redistributive.

We develop a two-period overlapping generations economy in a general equilibrium framework with endogenous capital formation. The economy is infinitely lived and the analysis is performed in steady state. Agents work in the first period and live in retirement in the second period. There are two types of agents who differ in terms of their productivity. The high productivity agents are considered as rich, whilst the low productivity agents are referred as poor. Redistribution in this context aims to bridge the inequality between the rich and the poor. There is a welfare maximising planner, with some degree of inequality aversion. The planner raises revenue by taxing the young rich and then redistributes the money to the poor - both young and old. The planner thus chooses two instruments, the tax rate to levy on the rich, and the proportion to allocate to the young poor - with the remainder going to the old poor. This chapter extends Chapter 3 by considering the redistributive instruments simultaneously and introducing an element of inequality aversion on behalf of the planner.

Abstracting from the new elements of the model, the results remain broadly in line with Chapter 3. The key analytical results can be summarised as:

1. Redistribution adversely affects capital formation. The higher the intergenerational redistribution component, the greater the crowding out;
2. An increase in the proportion of rich leads to a lower tax rate;
3. An improvement in productivity leads to lower tax rates;
4. Increased inequality aversion on the part of the planner results in higher tax rates.

The simulations results suggest that there are some parameter values where it is optimal to have no redistribution at all. The overall results can be summarised as:

1. An increase in the proportion of the rich or productivity leads to a lower tax rate and a tilting towards intergenerational redistribution. We

suggest that whilst the increase in proportion of the rich or productivity and the resulting fall in the tax rate could potentially result in dynamic inefficiency, the tilting towards intergenerational redistribution ensures the economy remains dynamically efficient.

2. Higher inequality aversion leads to higher taxes but also result in a shift towards intragenerational redistribution. The result is analogous to (1) above, in that the tilting towards intragenerational reverses some of the fall in capital induced by the increase in the tax rate.

3. As the discount factor falls and the agents become increasingly impatient, the tax rate remains broadly unchanged but there is a shift in emphasis towards intragenerational redistribution. This ensures that the fall resulting from the higher discounting is compensated for by the tilting towards intragenerational redistribution.

4. As the population gets younger, the tax rate falls infinitesimally and there is a shift towards intragenerational redistribution. The tilting counters the increase in capital resulting from the fall in the tax rate.

Our results thus suggest that the tax rate and the timing of the redistribution work so as to ensure capital does not increase or fall significantly.

This ensures the economy remains dynamically efficient at all times. When we put the model to test for a sample of 27 countries, we find that in all developing countries the planner will undertake redistribution even without inequality aversion. However, in the OECD countries, it is only when the planner gives a higher weight to the utility of the poor that redistribution will take place.

The chapter can be linked to the growing literature on the optimal form of pensions system to deal with an ageing population (World Bank, 1994; Gill et al., 2004). The current approach focuses on redistribution where pensions is used as a redistributive instrument as is income support. We consider which of the two approaches is given a higher weight depending on the prevailing economic conditions. Pensions in this set up are targeted in that only the poor benefit from such payments whilst the rich fund for their own retirement. The model can also be seen as incorporating two of the three pillars of the World Bank, firstly an element whereby everyone saves a proportion from their income; secondly, where pensions aim to ensure no one lives in poverty during retirement and redistributes from the rich to the poor.

In this chapter, we add to literature pertaining to inequality and redistribution. There is a large literature that emphasises the harmful impact of inequality on growth and how it can have serious politico-economic ramifications (Persson and Tabellini (1994); Alesina and Rodrik (1994); Alesina and Perrotti (1996)). Previously, redistribution had been considered as either intragenerational (Haveman, 1988) or intergenerational (Krueger and Kubler, 2006). In Chapter 3, we bridged this gap by comparing both types and considered which of the two tools would prevail under given economic conditions. In this paper, we include both intergenerational and intragenerational elements simultaneously. Our results can thus be considered as analogous to that of Conde-Ruiz and Galasso (2005) who suggest that with sufficient inequality in earnings and elderly in the economy, there is an equilibrium that supports the existence of both intra and intergenerational redistribution. In the current framework, we also find that both inter and intragenerational redistribution operate depending; with the weight attached to each depending on the economic environment.

The rest of this chapter is as follows: in Section 2 we describe the set up of the economy. Section 3 derives the competitive equilibrium whilst Section 4 considers the planner's problem in terms of finding the optimal level and timing of transfer. Section 5 tests the model to a sample of 27 countries.

Section 6 concludes and provides some potential extensions.

4.2 The Economy

We consider an infinitely lived overlapping generations economy in the Samuelson-Diamond spirit. Economic activity takes place over discrete time $t \in \{0, 1..., \infty\}$ without uncertainty. The economy consists of two types of utility maximising agents, $h(r,\ p)$; heterogeneous in their productivity ψ; profit maximising perfectly competitive firms; and a welfare-maximising social planner. We consider a planner who can potentially have an aversion to inequality such that the utility of the poor can be given a higher weight in the welfare function. At each time t, two factors, an amount of capital, k, and labour, l, are available as inputs to production and a homogeneous good, y, is produced.

4.2.1 Agents

At time t, one young generation and an old generation live simultaneously. Population grows at a constant rate n. Hence, at any time t, there are $(1+n)$ more (young) workers than (old) retirees. The population at t can thus be expressed as: $L_t + L_{t-1} = (2+n)L_t$. L_t refers to the agents born at t. Following

Samuelson (1958), the population structure is considered to be stationary. Therefore, the proportions and types of individuals remain the same across generations.

When young, each agent provides one unit of labour inelastically. When old the labour endowment is zero. There are two types of labour who differ in terms of their productivity $\psi \in (0, 1]$. In competitive markets, the productivity determines the wages the agents receive. We will term the high productivity agents as rich (r) and normalise their productivity to 1. Any agents with $\psi < 1$ will be considered as poor (p) and can potentially receive benefits. The rich and the poor make up proportions π and $(1 - \pi)$ of the economy respectively.

Agents maximise utility by maximising consumption $u^h(c_t^h, c_{t+1}^h)$ subject to the budget constraints over their lifetime. They are non-altruistic, they are born without any assets and leave no bequests. When young, the agents choose the level of consumption and savings which maximises their utility, whilst when old the agents live off their savings (and the poor potentially receive some transfers). For ease of manipulation, the intertemporal utility function is taken to be additive and log-linear. The utility function is thus strictly concave, since more consumption is preferred to less, and twice differentiable: $u'(c) > 0$ and $u''(c) < 0$. The function also satisfies $\lim_{c \to 0} u'(c) = \infty$ such that subject to its disposable income, the household will always choose a positive level of consumption when maximising life-cycle utility. We can express the utility-maximising problem of a rich agent as follows:

$$\begin{aligned} \underset{\{c_t^y, c_{t+1}^o, s_t\}}{Max} \quad &: \quad u = \ln c_t^{y,r} + \delta \ln c_{t+1}^{o,r} && (1) \\ subject\ to&: && (2) \\ c_t^{y,r} \quad &= \quad w_t (1-\tau) - s_t^r && (3) \\ c_{t+1}^{o,r} \quad &= \quad R_{t+1} s_t^r && (4) \end{aligned}$$

Eqn. (1) suggests that the agent maximises utility over both periods of his lifetime and future consumption is discounted by a factor $\delta \in (0, 1)$. The budget constraints suggest when young, the agent earns a wage w_t out of which a (constant) proportion τ is taxed. In line with Diamond (1965), the disposable income is then either consumed or saved. When old, the agent consumes his savings and the interest rates earned thereon. We note the rich fund their own retirement.

Similarly the utility function of the poor can be written in similar fashion with the difference that since the productivity of the poor is lower, their wage is lower as well. The poor receive a wage $\psi \in (0,1]$ times that of the rich such that their wage is ψw_t. We can express their problem as follows:

$$\begin{aligned} \underset{\{c_t^y, c_{t+1}^o, s_t\}}{Max} &: \quad u^p = \ln c_t^{y,p} + \delta \ln c_{t+1}^{o,p} \qquad (5) \\ subject\ to &: \\ c_t^{y,p} &= \psi w_t + b_t - s_t \qquad (6) \\ c_{t+1}^{o,p} &= R_{t+1} s_t + p_{t+1} \qquad (7) \end{aligned}$$

The utility maximising problem of the poor is similar to that of the rich in that it depends on consumption over their lifetime. The budget constraints of the poor suggest they receive a wage ψw_t and they can also receive benefits. Benefits in the first period is termed b_t whilst if it is in the second period it is considered as p_{t+1}. The poor agent allocates his first period disposable income between consumption and savings. The savings carried over plus the interest earned thereon funds the second period consumption along with any benefits the poor may receive in the form of pensions.

4.2.2 Firms

A large number of identical firms produce a homogeneous good using an identical economy-wide Cobb Douglas production function. The production function can be represented as $Y = K^{\alpha}L^{1-\alpha}$, where α is the share of capital in production.

In intensive form, the production function reduces to: $y = k^{\alpha}$. Firms maximise profit by taking factor prices, which are paid their marginal products in a competitive setting, as given. It is assumed that the labour market clears such that labour demand equals labour supply and the wages received by a worker depends on his level of productivity. The economy is endowed with an initial capital stock $K_0 > 0$ and capital depreciates fully from one period to the next[10]. We assume no technological change. The production function satisfies the usual conditions such that $f(0) = 0$, $f'(k) > 0$, $f''(k) < 0$ and the Inada conditions: $\lim_{k \to 0} f'(k) = \infty$ and $\lim_{k \to \infty} f'(k) = 0$.

4.2.3 Social Planner

The social planner is considered as a social welfare maximiser and we allow for it to have an aversion to inequality ϵ. Its policy role is limited to that of raising taxes from the rich at a rate τ and deciding on the optimal proportion ω to transfer to the young poor and the remainder $(1 - \omega)$ to the old poor. The total tax raised (T) is allocated between benefits to the young (B)

[10] Depreciation is used in the non-conventional sense: the old consume the capital and the young replenish it.

and benefits to the old (P). Hence, $T_t = \pi L_t \tau w_t = B_t + P_t$. B_t refers to the component of taxes redistributed in the form of income support and it is allocated to the proportion of poor born at time t. P_t refers to the component of redistribution that takes place in the form of pensions and it is allocated to the proportion of poor born at time $t-1$. We can then represent the equilibrium conditions as follows:

$$\begin{aligned} B_t &= \omega T_t = (1-\pi) L_t b_t \\ P_t &= (1-\omega) T_t = (1-\pi) L_{t-1} p_t \end{aligned}$$

With a stationary population structure and commitment device, the poor are thus faced with the following benefits levels in intensive form:

$$b_t = \frac{\omega \pi L_t \tau w_t}{(1-\pi) L_t} = \frac{\omega \pi \tau w_t}{(1-\pi)} \tag{8}$$

$$p_t = \frac{(1-\omega) \pi L_t \tau w_t}{(1-\pi) L_{t-1}} = \frac{(1-\omega) \pi (1+n) \tau w_t}{(1-\pi)} \tag{9}$$

If $\omega = 1$, all the tax raised is redistributed to the young, in which case we have a fully funded pensions scheme, since all agents provide for their own pensions. On the other hand, if $\omega = 0$, all the taxes raised are allocated to the old in the form of pensions.

4.3 Competitive Equilibrium

Given the households' and the firms' objectives, a competitive equilibrium for the economy can be defined as a sequence of consumption $\{c_t^y, c_t^o\}_{t=0}^{\infty}$ such that:

1. Given a sequence of taxes and transfers, $\{\tau w_t, b_t/p_{t+1}\}_{t=0}^{\infty}$, and the prevailing competitive wages, w_t, and interest rate, r_t, solves the individual's optimisation problem subject to satisfying the Euler equation;

2. Factors of production are paid their marginal products ($w_t = (1-\alpha)k_t^{\alpha}; R_t = \alpha k_t^{\alpha-1}$) and labour and capital markets clear such that $L_t^D = L_t$ and $S_t = K_{t+1}$;

3. The planner's budget is always balanced. Hence taxes raised is redistributed between the young and old poor in the same period: $T_t = B_t + P_t$;

4. The economy's resource constraint is always satisfied. In intensive form, the constraint is defined as the allocation of current output, y_t, between consumption of the young and old and capital formation in the next period.

$$y_t = c_t^y + \frac{c_t^o}{(1+n)} + (1+n)k_{t+1} \tag{10}$$

Given the above definition of competitive equilibrium, we can express the intertemporal budget constraints for the rich and the poor as:

$$\begin{aligned} c_t^{y,r} + \frac{c_{t+1}^{o,r}}{R_{t+1}} &= w_t(1-\tau) && (11) \\ c_t^{y,p} + \frac{c_{t+1}^{o,p}}{R_{t+1}} &= \psi w_t + b_t + \frac{p_{t+1}}{R_{t+1}} && (12) \end{aligned}$$

We can then write the Lagrangian and derive the standard Euler equation which dictates the optimal allocation of consumption over the two periods of an agent's lifetime. For the poor, this can be expressed as:

$$c_{t+1}^o = \delta R_{t+1} c_t^y \tag{13}$$

We can now proceed to derive the optimal levels of consumption and savings of both agents as:

$$\begin{aligned} c_t^{y,r} &= \left(\frac{[w_t(1-\tau)]}{1+\delta}\right) && (14) \\ c_{t+1}^{o,r} &= \left(\frac{\delta R_{t+1}[w_t(1-\tau)]}{1+\delta}\right) && (15) \\ s_t^{y,r} &= \left(\frac{\delta[w_t(1-\tau)]}{1+\delta}\right) && (16) \end{aligned}$$

$$
\begin{aligned}
c^{y,p} &= \left(\frac{1}{1+\delta}\left[\left(\psi+\frac{\omega\pi\tau}{(1-\pi)}\right)w_t+\frac{(1-\omega)\,\pi\,(1+n)\,\tau w_{t+1}}{(1-\pi)\,R_{t+1}}\right]\right) && (17)\\
c^{o,p} &= \left(\frac{\delta R_{t+1}}{1+\delta}\left[\left(\psi+\frac{\omega\pi\tau}{(1-\pi)}\right)w_t+\frac{(1-\omega)\,\pi\,(1+n)\,\tau w_{t+1}}{(1-\pi)\,R_{t+1}}\right]\right) && (18)\\
s_t^{y,p} &= \left(\frac{1}{1+\delta}\left[\delta\left(\psi+\frac{\omega\pi\tau}{(1-\pi)}\right)w_t-\frac{(1-\omega)\,\pi\,(1+n)\,\tau w_{t+1}}{(1-\pi)\,R_{t+1}}\right]\right) && (19)
\end{aligned}
$$

Eqns. (14)-(16) show that the equilibrium levels of consumption and savings of the rich depends on their disposable income. On the other hand, eqns (17)-(19) suggest that for the poor, their consumption levels depend on their income plus any income support they receive and the discounted value of pensions. We also note from eqn. (19) that there is potential crowding out that arises when pensions are provided. If $\omega < 1$, then this reduces the amount of savings undertaken by the agents. Maximum crowding out occurs when $\omega = 0$, i.e., when redistribution is only through pensions.

We can now consider how capital formation takes place in this economy. With full depreciation of capital from one period to the next, the capital in the next period is simply the savings of the current working (young) generation. This can be characterised as:

$$
K_{t+1} = S_t = S_t^r + S_t^p \qquad (20)
$$

suggesting that capital in the next period consists of the savings of the working rich and poor. In intensive form, this reduces to:

$$(1+n)\,k_{t+1} = \pi s_t^r + (1-\pi)\,s_t^p$$

Since the savings of the rich and poor differs, the proportions of rich and poor play an important role in the capital accumulation process. In steady state where $k_{t+1} = k_t = k^*$, steady-state capital is characterised as:

$$k^* = \left[\left(\frac{\delta\alpha(1-\alpha)\left\{\pi\,(1-\tau)+(1-\pi)\,\psi+\omega\pi\tau\right\}}{(1+n)\left[(1+\delta)\,\alpha+(1-\omega)\,\pi(1-\alpha)\tau\right]}\right)\right]^{\frac{1}{1-\alpha}} \tag{21}$$

Proposition 9 *(a) An increase in ω (intragenerational transfers) increases k*

(b) An increase in τ decreases k

Proof. The Proof is in Appendix 4.1 ∎

From eqn.(21), it follows that if $\omega = 1$, then the tax rate does not affect capital accumulation if all redistribution is done in the first period. Steady state capital is given by:

$$k^* = \left[\left(\frac{\delta(1-\alpha)\left\{\pi+(1-\pi)\,\psi\right\}}{(1+n)\,(1+\delta)}\right)\right]^{\frac{1}{1-\alpha}} \tag{22}$$

Proposition 10 *(a) An increase in δ increases k*

(b) An increase in π increases k

(c) An increase in ψ increases k

Proof. The Proof is in Appendix 4.1 ■

If $\omega = 0$, i.e., implying all the redistribution is intertemporal, then the tax rate matters:

$$k^* = \left[\left(\frac{\delta\alpha(1-\alpha)\left\{\pi\left(1-\tau\right)+\left(1-\pi\right)\psi\right\}}{\left(1+n\right)\left[\left(1+\delta\right)\alpha+\pi(1-\alpha)\tau\right]} \right) \right]^{\frac{1}{1-\alpha}} \tag{23}$$

It thus follows from proposition 1(a) that capital is greater for (22) than (23).

4.4 The Planner's Problem

The planner's objective is to maximise aggregate social welfare. Given this objective and the competitive equilibrium, the planner's problem is to choose τ and ω so as to maximise $\{c_t^y, c_t^o, k_{t+1}\}_{t=0}^{\infty}$ subject to the allocation $\{y_t\}_{t=0}^{\infty}$. We also account for the fact that the planner can have an aversion to inequality such that the utility of the poor is given a higher weight $\epsilon \geqslant 1$. If $\epsilon = 1$, the welfare function is Benthamite where all agents are given the same weight. Assuming steady state, the planner's problem can be expressed as follows:

$$\underset{\{\tau,\omega\}}{Max} : V = \sum_{t=0}^{\infty} \rho^t \left(U\left(c^r\right) + \epsilon U\left(c^p\right)\right) \tag{24}$$

The welfare function consists of the utility of the rich and poor and the inequality aversion parameter of the planner, $\epsilon \geqslant 1$. ρ refers to the weight the planner attaches to the generations, living and unborn, and this is decreasing in time.

Given the additive nature of the utility functions of the agents, we can proceed to write the planner's problem as that of maximising the utility of all agents living simultaneously, that is, the young and old, rich and poor. We can then write this in intensive form as:

$$V = \pi \left[(1+n)\ln(c^{y,r}) + \ln(c^{o,r})\right] + \epsilon\,(1-\pi)\left[(1+n)\ln(c^{y,p}) + \ln(c^{o,p})\right] \tag{25}$$

The first part of the welfare function refers to the proportion of rich in the economy. Since population grows at a rate n, there are $(1+n)$ more youngsters than elderly. The second part of the welfare function considers the poor and ϵ is the inequality aversion parameter of the planner. The higher the value of ϵ, the greater the inequality aversion.

4.4.1 The Optimal Level of τ & ω

We proceed to differentiate V with respect to τ and ω to find their respective optimal levels. For the general case, this results in:

$$
\begin{aligned}
\frac{\partial V}{\partial \tau} &= \frac{-\pi}{(1-\tau)} + \left[\frac{[(3+n)\alpha-1]\pi+[(1+n)+\alpha]\epsilon(1-\pi)}{(2+n)(1-\alpha)}\right]\left[\frac{(\omega-1)\pi}{\{\pi(1-\tau)+\omega\pi\tau+(1-\pi)\psi\}} - \frac{(1-\omega)\pi(1-\alpha)}{[(1+\delta)\alpha+(1-\omega)\pi(1-\alpha)\tau]}\right] \\
&+\epsilon(1-\pi)\left[\begin{array}{c} \frac{(1-\pi)\psi(1-\omega)\pi(1-\alpha)+(1+\delta)\alpha\omega\pi+2\omega\pi(1-\omega)\pi(1-\alpha)\tau+\delta(1-\alpha)(1-\omega)\pi[\pi(1-2\tau)+2\omega\pi\tau+(1-\pi)\psi]}{[[(1-\pi)\psi+\omega\pi\tau][(1+\delta)\alpha+(1-\omega)\pi(1-\alpha)\tau]+\delta(1-\alpha)\{\pi(1-\tau)+\omega\pi\tau+(1-\pi)\psi\}(1-\omega)\pi\tau]} \\ \\ -\frac{(\omega-1)\pi}{\{\pi(1-\tau)+\omega\pi\tau+(1-\pi)\psi\}} \end{array}\right] \\
\frac{\partial V}{\partial \omega} &= \left[\frac{[(3+n)\alpha-1]\pi+[(1+n)+\alpha]\epsilon(1-\pi)}{(2+n)(1-\alpha)}\right]\left[\frac{\pi\tau}{\{\pi(1-\tau)+\omega\pi\tau+(1-\pi)\psi\}} + \frac{(1-\alpha)\pi\tau}{[(1+\delta)\alpha+(1-\omega)\pi(1-\alpha)\tau]}\right] \\
&+\epsilon(1-\pi)\left[\begin{array}{c} \frac{-(1-\pi)\psi(1-\alpha)\pi\tau+(1+\delta)\alpha\pi\tau+(1-2\omega)(1-\alpha)\pi\tau\pi\tau-\delta(1-\alpha)\pi\tau\{\pi(1-\tau)+(1-\pi)\psi\}+\delta(1-\alpha)(1-2\omega)\pi\tau\pi\tau}{[[(1-\pi)\psi+\omega\pi\tau][(1+\delta)\alpha+(1-\omega)\pi(1-\alpha)\tau]+\delta(1-\alpha)(1-\omega)\pi\tau\{\pi(1-\tau)+\omega\pi\tau+(1-\pi)\psi\}]} \\ \\ -\frac{\pi\tau}{\{\pi(1-\tau)+\omega\pi\tau+(1-\pi)\psi\}} \end{array}\right]
\end{aligned}
$$

Given the nature of the two simultaneous equations, we have recourse to simulations to solve for τ and ω in the next section. For now, let us assume that $\omega = 1$, such that all redistribution is intragenerational. Then, $\frac{\partial V}{\partial \tau} = \frac{-1}{(1-\tau)} + \left[\frac{\epsilon(1-\pi)}{[(1-\pi)\psi+\pi\tau]}\right]$ and it follows that:

$$\tau = \frac{(1-\pi)(\epsilon-\psi)}{\pi+\epsilon(1-\pi)} \tag{26}$$

Proposition 11 *(a) An increase in π reduces τ*

(b) An increase in ψ reduces τ

(c) An increase in ϵ increases τ

Proof. The Proof is in Appendix 4.1 ■

The results are consistent with expectations in that the richer an economy or the lower the inequality, the less the need for redistribution. Similarly, the higher the inequality aversion of the planner, the higher the tax rate. When we set $\epsilon = 1$, the optimal tax rate is the same as the one for intragenerational transfer in Chapter 3 and yields $\tau = (1 - \pi)(1 - \psi)$. This is consistent with the view that complete equalisation of consumption of the rich and poor is optimal.

4.4.2 Simulation Results and Sensitivity Analysis

We now consider the behaviour of the model for a given set of parameters. These parameters have been chosen to initially yield a dynamically efficient solution and provide plausible values of τ and $\omega \in (0, 1)$. There exists a set of parameters for which redistribution is not optimal. We ignore these situations. We assume the agents have the same discount factor which is taken to be line in line with values used for macro simulations (see de la Croix and Michel, 2002; Krueger and Kubler, 2006). We set the discount factor at 0.96. The population growth rate is set at 1 percent. In the baseline, we assume that the poor earn 75 percent of what the rich earn and we assume that the rich, π, make up 40 percent of the economy, with the poor

	α	δ	ϵ	π	ψ	n			
	0.3	0.96	1	0.4	0.75	0.01			
	Capital	**Output**	**Wage**	**Interest**	$\boldsymbol{\tau}$	$\boldsymbol{\omega}$	UR	UP	**V**
V_0	0.169	0.587	0.411	1.04	0	0	-3.063	-3.628	-3.489
V_1	0.156	0.572	0.401	1.10	0.1521	0.5898	-3.380	-3.380	-3.466

Table 4: Baseline Simulations

accounting for the remaining 60 percent. We initially set $\epsilon = 1$ whereby the planner treats all agents equally. We allow for all these values to change when carrying out sensitivity analysis.

Baseline Results The baseline results are presented in Table 4:

Two welfare functions are presented in Table 4. For comparative purposes, V_0, which represents the case where there is no redistribution, has been included. V_1 represents the welfare function with redistribution. We allow for redistribution to take place over the two periods of the lifetime of the agent. In V_0, the utility of the agents is only a function of their earnings and we can find that as expected, the utility of the rich is higher than that of the poor.

After allowing for redistribution to take place, we note that the optimal tax rate on the rich is around 15.21 percent and the young poor receive around 59 percent of the tax raised in the form of income support whilst the old poor receive the remaining 41 percent in the form of pensions. As a result of the redistribution, we note that capital is lower and as a result output and wages are lower too. This results in a higher interest rate. Though before redistribution the utility of the rich was higher than that of the poor, following redistribution, the utility of the rich falls unambiguously whilst that of the poor increases and they are both equalised. Aggregate welfare is unambiguously higher.

The results differ from the situation with one instrument (Chapter 3) whereby, when the agents had the same discount factor, capital formation was not affected. However, it relied on the fact that all the redistribution was intergenerational. In this case, since we have both inter and intragenerational redistribution simultaneously, capital is lower and as a result we have a higher interest rate.

	k	**y**	**w**	**R**	**Tax**	ω	UR	UP	Welfare
α	$-$	$-$	$-$	$+$	$-$	$+$	$-$	$-$	$-$
δ	$+$	$+$	$+$	$-$	$+$	$-$	$-$	$-$	$-$
ϵ	0	0	0	0	$+$	$+$	$-$	$+$	$-$
π	$+$	$+$	$+$	$-$	$-$	$-$	$+$	$+$	$+$
ψ	$+$	$+$	$+$	$-$	$-$	$-$	$+$	$+$	$+$
n	$-$	$-$	$-$	$+$	$-$	$+$	$-$	$-$	$-$

Table 5: Impact of a 1 percent change

We now consider the impact of a 1% increase in the exogenous parameters on the endogenous variables relative to the baseline results. These are summarised in Table 5:

The results can be summarised as follows: Since there is a positive relationship between capital and both output and wages, any parameter that leads to an increase in k, leads to an increase in output and wages. Conversely, an increase in capital leads to a fall in the interest rate.

α and n are negatively related with capital and have exactly the same impact on the parameters under consideration. A increase in n, which results in a younger population, leads to a decrease in the tax rate. Moreover, the increase in n also results in a tilting towards intragenerational redistribution. This leads to an unambiguous decline in welfare of all the agents concerned. We can consider the tax rate as an "equilibrating mechanism". An increase in n results in a lower steady state capital. The fall in the tax rate and the tilting towards intragenerational redistribution helps to restore k such that some of the initial fall is reversed.

Increases in δ, π and ψ lead to increases capital . However, the impact differs so far as taxation is concerned. An increase in δ leads to an infinitesimal increase in the tax rate. However, increases in both π and ψ lead to declines in the tax rate. The latter could be consistent with the fact that increases in π and ψ broaden the tax base and as a result the optimal tax rate falls. Increases in ψ reduce the productivity gap between the poor and the rich and this also implies a lower amount of taxation is required. Moreover, as mentioned earlier, the increase in π and ψ reduce poverty and inequality in the economy and as such the need for redistribution falls. Increases in δ, π and ψ tilt redistribution towards intergenerational, in that ω falls. This is consistent with our previous suggestion of the "equilibrating mechanism" at work and the tilting ensures the economy remains dynamically efficient. Whilst a marginal 1 percent increase in δ leads to a reduction in welfare (possibly due to the fact that the utility lost out of reduced consumption exceeds the gain from the increased capital) increases in π and ψ unambiguously increase welfare.

Sensitivity Analysis The baseline results gave an overview of the optimal level of tax rate and ω for the given set of parameters. However, to get a better understanding of the behaviour of the endogenous variables, we now allow for the exogenous parameters to vary. This is done with the aim of getting a better understanding of the tilting the policy variables with respect to changes in the selected parameters. We have to note that redistribution only takes place for a given range of parameters, outside which it is no longer optimal to redistribute.

Proportion of Rich The results on the proportion of rich are as before. An increase in the proportion of rich has a positive impact on capital and hence results in higher output and wages with a resulting lower interest rate.

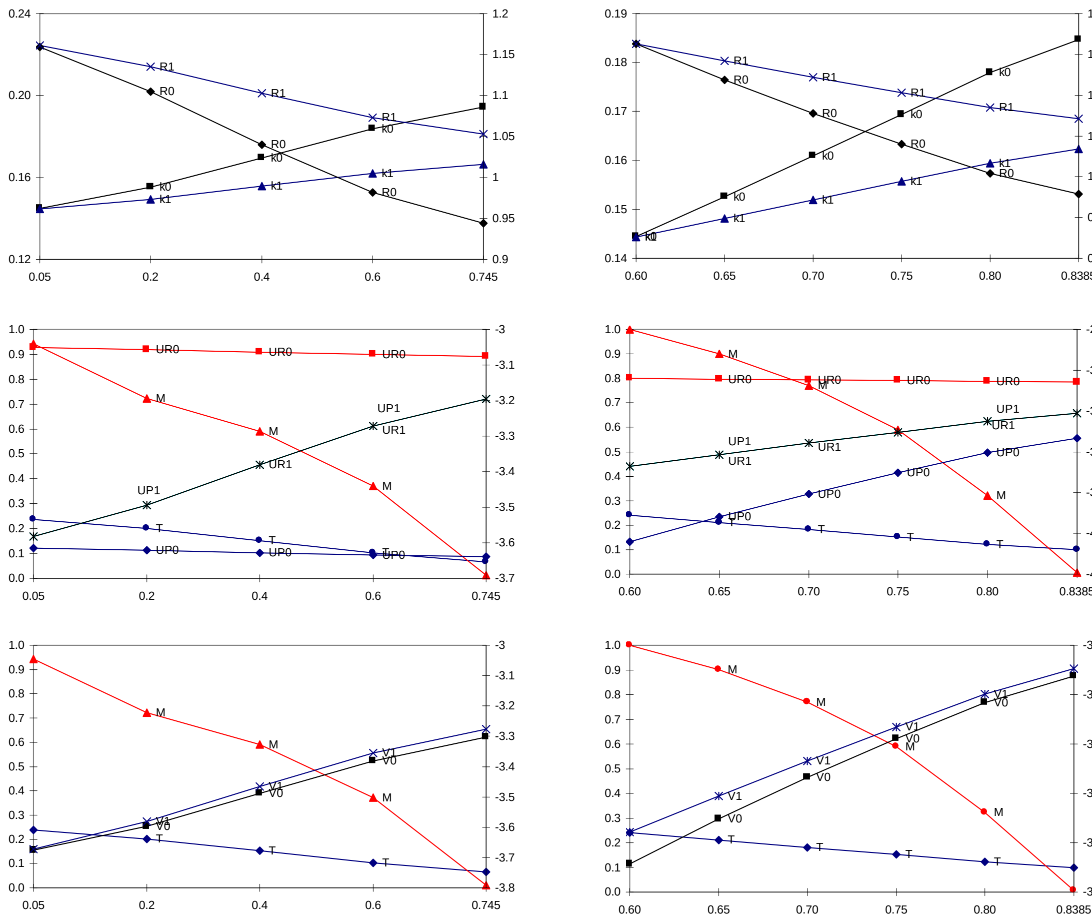


Figure 11: Impact of changes in Pi (L) and Psi (R)

However, we also note that as the proportion of rich increases, the tax rate falls. For instance, when the rich represent only 5 percent of the population, the tax rate is around 24 percent and this falls to 6.57 percent when the rich account for 74.5 percent of the population (which is near to the upper bound for redistribution). It can also be seen that as the proportion of rich increases, the distribution tilts towards intergenerational redistribution. When π is 5 percent, around 94 percent of the redistribution is intragenerational, but it falls to only 1 percent when π is 74.5 percent. Consistent with our suggestion that the optimal tax rate and timing of redistribution act as an equilibrating device, the tilting in the redistribution between generations leads to a relatively stable capital stock such that the overall impact on the interest rate is more subdued. The economy thus remains dynamically efficient.

Productivity of Poor The same mechanism applies for the productivity parameter. As the productivity of the poor increases, the same effects are noted on capital. The tax rate falls as well - from 24 percent, when the productivity of the poor is 60 percent that of the rich, to around 10 percent, when the productivity of the poor rises to 84 percent that of the rich. ω tilts towards intergenerational redistribution. When ψ is 0.6, all redistribution is intragenerational, when it rises to 0.84 all redistribution is intergenerational. Welfare improves unambiguously.

The results are consistent with the suggestion that as the productivity of the poor increases, this reduces the gap between the rich and the poor thereby resulting in a fall in the tax rate. The fall in the tax rate leads to an increase in capital but the tilting in ω reverses some of the increase in capital and ensures the economy remains dynamically efficient. We can see in Figure 11 that without redistribution, there is potential for the economy to become dynamically efficient.

Inequality Aversion We initially assumed that the planner attached the same weight to all agents. We now allow for the fact that the planner can be averse to inequality (for whatever reason) and attaches a higher weight to the welfare of the poor. The results are summarised in Figure 12. The first thing we note is that as the planner attaches a higher weight to the utility of the poor, this leads to an increase in the tax rate which rises from 15.21 percent when the planner attaches the same weight to all agents to 26.27 percent when he gives an extra 25 percent weight to the utility of

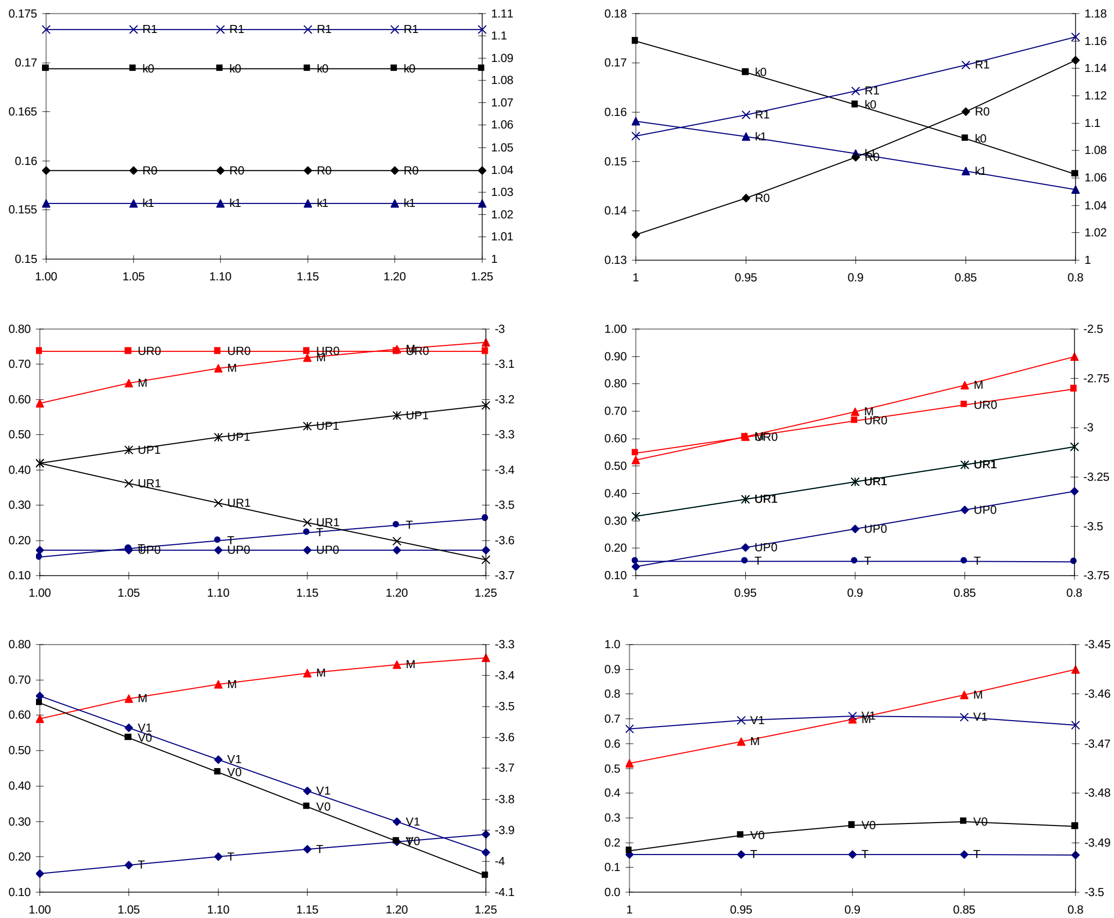


Figure 12: Impact of changes in Epsilon (L) and Delta(R)

the poor. There is also a tilting in redistribution between generations in that as the inequality aversion of the planner increases, this leads to an increased emphasis on intragenerational redistribution. However, there is an interesting result in so far as the overall impact on capital is concerned. Even though increases in inequality aversion lead to a higher tax rate, this does not affect overall capital formation and as a result leaves output, wages and the interest rate unchanged. We suggest this is due to the fact that the higher tax rate is countered by the shift towards intragenerational redistribution which mitigates the negative impact on capital. The utility of the rich and the poor is no longer equalised in that as the planner's inequality aversion increases, the welfare of the rich falls while that of the poor increase unambiguously. Aggregate welfare is lower as a result of higher inequality aversion.

Discount Factor We initially assume agents do not discount the future and then allow for the discounting of the future at an increasing rate. The overall results are in line with expectations and are summarised in Figure 12. As δ falls, this leads to a lowering of capital resulting in the incipient falls in output and wages and a fall in the interest rate. The tax rate remains in the region of 15 percent in that as δ falls from 1 to 0.8, the tax rate only falls by 14 basis points, from 15.22 percent to 15.08 percent. However, there is a significant tilt in ω. For the same range of δ, ω increases from 0.52 to 0.899. This means that as the agents attach a higher weight to consumption in the first period, an increasing weight is attached to intragenerational redistribution. Thus, an increasing proportion of the benefits is given to the poor agents when they are young. However, this mechanism can be rationalised on the grounds that intragenerational redistribution has less of an adverse impact on capital formation as opposed to intergenerational redistribution. In this case, the fall in δ is counteracted by the increase in ω such that the overall impact on capital is reduced, though capital ends up lower. The impact of δ on aggregate welfare suggests that as δ initially falls, welfare increases, however, beyond a certain point, this reaches a turning point and as δ falls further, this is accompanied by a fall in welfare as well. This is due to fact that, as consumption initially increases, it adds to welfare and beyond a certain point, this gain in welfare is overturned by the adverse impact on capital formation.

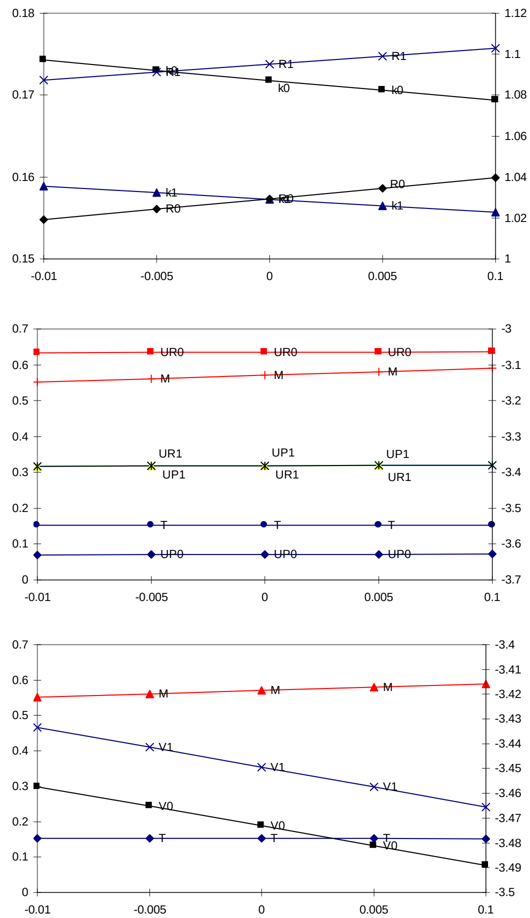


Figure 13: Impact of Change in Population Growth Rate

Population Growth Rate An increase in the population growth rate reduces capital, output and wages whilst it leads to an increase in the interest rate. Though there is an infinitesimal fall in the tax rate, redistribution shifts towards intragenerational, albeit at a slower pace. Aggregate welfare is unambiguously lower in part due to the fall in capital. The fall in the tax rate and the increased emphasis on intragenerational transfers helps to reverse some of the fall in capital.

4.5 An Application

We now consider the applicability of the model to a sample of 27 countries made up as follows: Europe(12), America(7), Africa(4), Asia(2), Pacific(2). We use the same estimates of $\alpha, \delta, \pi, \psi$, and n as previously. The value of ϵ is then fixed to get estimates of τ and μ for each of the countries. As before, we assume that the discount factors of the two types of agents are the same and the only element of heterogeneity is in productivity. Data sources and definitions are provided in Appendix 3.2. An overview of the endogenous variables is given in the Table 6 and the results are summarised in Table 7.

Country	n	Alpha	Beta	Pi	Psi	R
Arg	1.3	0.26	0.91	0.76	0.50	1.11
Auz	1.3	0.30	0.93	0.75	0.80	1.06
Aus	0.4	0.26	0.95	0.66	0.75	1.05
Blg	0.2	0.26	0.93	0.73	0.63	1.08
Brz	1.7	0.25	0.65	0.58	0.38	1.36
Can	1.1	0.23	0.95	0.75	0.67	1.07
Chl	1.5	0.25	0.90	0.63	0.47	1.16
Egy	2.0	0.25	0.95	0.51	0.47	1.19
Fin	0.4	0.32	0.96	0.68	0.85	1.04
Fra	0.5	0.23	0.95	0.71	0.65	1.08
Ger	0.2	0.27	0.92	0.66	0.53	1.12
Ire	0.8	0.22	0.97	0.65	0.75	1.06
Isr	2.3	0.29	0.86	0.75	0.50	1.13
Ita	0.1	0.27	0.95	0.64	0.95	1.03
Mex	1.6	0.21	0.95	0.58	0.49	1.15
Neth	0.6	0.28	0.95	0.73	0.75	1.06
NZ	1.1	0.29	0.93	0.69	0.74	1.07
Nig	2.7	0.25	0.96	0.41	0.35	1.29
Per	1.8	0.29	0.88	0.76	0.62	1.10
Sng	2.3	0.22	0.95	0.74	0.45	1.11
SA	2.0	0.27	0.96	0.64	0.50	1.13
Spn	0.6	0.22	0.96	0.64	0.64	1.09
Swi	0.6	0.25	0.97	0.72	0.71	1.06
Tur	1.8	0.24	0.96	0.43	0.42	1.24
Ugn	3.2	0.17	0.96	0.23	0.35	1.39
UK	0.3	0.22	0.96	0.75	0.72	1.05
US	1.0	0.27	0.95	0.77	0.65	1.07

Table 6: Parameters Summary

	E=1		**E=1.1**		**E=1.25**		**E=1.5**		**E=2**	
Country	Tau	Mu								
Brz	0.27	0.69								
Chl	0.21	0.27								
Egy	0.27	0.42								
Fin	0.05	0.64								
Ger	0.17	0.32								
Isr	0.13	0.60								
Nig	0.39	0.74								
Per	0.10	0.30								
SA	0.19	0.40								
Tur	0.34	0.50								
Ugn	0.53	0.13								
Arg			0.15	0.10						
Auz			0.08	0.15						
NZ			0.11	0.22						
Blg					0.16	0.07				
Mex					0.31	0.11				
Neth					0.13	0.10				
US					0.14	0.04				
Aus							0.22	0.23		
Sng							0.25	0.12		
Spn							0.27	0.02		
Swi							0.20	0.04		
Can									0.27	0.17
Fra									0.31	0.26
Ire									0.33	0.14
Ita									0.28	0.33
UK									0.26	0.03

Table 7: Inequality Aversion and Redistribution

As mentioned earlier, for some range of parameter values and assuming the planner treats everyone equally so there is no inequality-aversion, it is not optimal to have redistribution from the rich to the poor. However, as inequality aversion is introduced, there is an optimal level of tax that is levied on the rich and this is then redistributed to the poor. In Table 7, as the value of ϵ increases, so does the inequality aversion of the planner. We find that in the first column, all the agents are treated equally and in that case it is optimal to redistribute in only 11 of the 27 countries. As inequality aversion increases, more countries redistribute. For instance for an inequality aversion of an extra 10 percent, 3 more countries redistribute and with $\epsilon = 2$ all the 27 countries redistribute.

Some of the results can be interpreted as follows: We can see that all the developing countries in the sample are lumped in the first column - and these are the countries where, on average, inequality and the proportion of poor is highest. Hence, we find that even without having an inequality aversion, the planner redistributes. The tax rates vary from 5 percent to 53 percent increasing in line with inequality and the proportion of the poor. The value of ω varies between a minimum of 13 percent and a high of 74 percent suggesting that for some countries a higher proportion of the redistribution is intergenerational, whilst in others in it is intragenerational. As we allow for inequality aversion to increase, more of the European (OECD) countries find it optimal to redistribute.

One of the main differences in the results this time round is the fact that redistribution takes place over both periods and the temporal element varies from one country to the other. We also find that whilst previously, after adjusting for α, it was optimal for all the countries to have intragenerational redistribution, with the introduction of inequality aversion the results are mixed. We find that in most cases when redistribution occurs as a result of inequality aversion, intergenerational redistribution accounts for the highest proportion.

4.6 Conclusions

We have considered what is the optimal timing of transfers from the rich to the poor. We assume that the planner has the option of giving a proportion to the agent when he is young and the remainder in the form of pensions in retirement. The question that we have addressed is what proportion should be allocated to the agents in which period of their lifetime and how does this vary with changes in the economic environment. Our basic results confirm that redistribution affects capital formation and the higher the intergenerational component of redistribution the greater the crowding out. However, redistribution is welfare enhancing in that the benefit for the poor more than exceeds the cost to the rich. Whilst the poor benefit from higher taxes and hence more redistribution, this leads to a decline in the welfare of the rich. We find that the higher the inequality and the higher the inequality aversion of the planner, the higher the tax rate. However, the increase in inequality through a reduction in the proportion of the rich or a fall in productivity and the increase in inequality aversion also leads to a tilting towards intragenerational redistribution. We attribute this to the fact that the increase in the tax rate leads to a fall in capital but the tilting in the timing of the redistribution leads to a reversal in the fall in capital. In richer economies, the tax rate falls but there is a switch towards intergenerational redistribution which ensures the economy remains dynamically efficient. Increased myopia leads to an infinitesimal change in the tax rate but favours intragenerational redistribution. Population ageing favours intergenerational redistribution. Our results suggest that redistribution in the economy acts like an equilibrating mechanism which ensures that falls (increases) in capital as a result of changes in the economic environment are reversed through a tilting in the timing of redistribution.

Appendix 4.1

Proof of Proposition 9:

$$(1a) : \frac{\partial k}{\partial \omega} = \frac{1}{1-\alpha}\left[\frac{[(1+\delta)\alpha+[\delta\alpha(1-\alpha)\{\pi(1-\tau)+(1-\pi)\psi\}+\pi\tau]\pi(1-\alpha)\tau]}{(1+n)[[(1+\delta)\alpha+(1-\omega)\pi(1-\alpha)\tau]]^2}\right]^{\frac{\alpha}{1-\alpha}} > 0$$

$$(1b) : \frac{\partial k}{\partial \tau} = \frac{1}{1-\alpha}\left[-\frac{\alpha(1-\alpha)(1-\omega)\pi\delta[\alpha(1+\delta)+(1-\alpha)\{\pi(1-\tau)+(1-\pi)\psi\}]}{(1+n)[[(1+\delta)\alpha+(1-\omega)\pi(1-\alpha)\tau]]^2}\right]^{\frac{\alpha}{1-\alpha}} < 0$$

Proof of Proposition 10

$$(2a) \quad : \quad \frac{\partial k}{\partial \delta} = \frac{1}{1-\alpha}\left[\frac{(1-\alpha)\{\pi+(1-\pi)\psi\}}{(1+n)(1+\delta)^2}\right]^{\frac{\alpha}{1-\alpha}} > 0$$

$$(2b) \quad : \quad \frac{\partial k}{\partial \pi} = \frac{1}{1-\alpha}\left[\frac{\delta(1-\alpha)(1-\psi)}{(1+n)(1+\delta)}\right]^{\frac{\alpha}{1-\alpha}} > 0$$

$$(2c) \quad : \quad \frac{\partial k}{\partial \psi} = \frac{1}{1-\alpha}\left[\frac{\delta(1-\alpha)(1-\pi)}{[(1+n)(1+\delta)]}\right]^{\frac{\alpha}{1-\alpha}} > 0$$

Proof of Proposition 11

$$(3a) \quad : \quad \frac{\partial \tau}{\partial \pi} = \frac{-(\epsilon-\psi)}{[\pi+\epsilon(1-\pi)]^2} < 0$$

$$(3b) \quad : \quad \frac{\partial \tau}{\partial \psi} = \frac{-(1-\pi)[\pi+\epsilon(1-\pi)]}{[\pi+\epsilon(1-\pi)]^2} < 0$$

$$(3c) \quad : \quad \frac{\partial \tau}{\partial \epsilon} = \frac{(1-\pi)[\pi+\psi(1-\pi)]}{[\pi+\epsilon(1-\pi)]^2} > 0$$

5 THE THREE PILLARS OF PENSIONS AND WELFARE

5.1 Introduction

With the ageing of the population of many countries, developed and developing, the issue of pensions reforms has attracted significant academic and policy interest. This was given a further impetus after the wave of pensions reforms that swept through Latin America, starting with Chile, in the early 1980s. The initial debate focused on a shift from the traditional PAYG pensions system to a fully-funded scheme, which would in theory be more suited to population ageing. However, since the publication of the World Bank's (WB) flagship report "Averting the Old Age Crisis" (1994), the general consensus has been that the best way to provide for pensions is through a multi-pillar system. The optimal design of pensions system has been in the limelight ever since. The WB suggested that any pensions scheme had to achieve 3 objectives and this was the main reason why it proposed the 3-pillars of pensions. Each of the pillars had one objective to fulfill:

1. The first pillar was compulsory and was set with the aim of ensuring

everyone had a basic level of pensions in retirement. This pillar would be publicly managed.

2. The second pillar, privately managed, would ensure that everyone is saving for retirement.

3. The third pillar was set to encourage an element of voluntary savings that would complement the income from Pillars 1 and 2.

The first pillar provided a basic level of pensions to the elderly irrespective of their level of contribution. In this respect, it had a redistributive element. On the other hand, the second pillar would provide pensions according to the level of contribution of the agents. The third pillar was considered as a residual whereby people would undertake savings for retirement such that their old-age consumption would be complemented with the pensions they receive under Pillars 1 and 2. "Averting the Old Age Crisis" suggested that by diversifying the pensions scheme, the risks would be more diversified and at the same time there are other economic benefits that could emerge. The 3-pillar scheme would be more efficient in that each pillar had clearly defined objectives that it was meant to satisfy. In turn, the fact that the benefits under pillars 2 and 3 is closely related to the level of contribution of the individuals would reduce distortions in the economy. There were also other benefits that the 3-pillars were meant to achieve in terms of increasing long term saving, deepening capital markets and promoting growth. Such a system would enable a better diversification of not only economic risks but also political risks.

However, the WB Report also acknowledged that the three pillar scheme was best suited to the formal sector and in many developing countries, where the pensions schemes are not well developed, if existing at all, the informal sector had a key role to play. As such, everyone might not be covered under the 3 pillars. In this respect, the WB added 2 further pillars to its scheme. This was published in another influential WB Report ("Old-Age Income Support in the Twenty-first Century: An International Perspective on Pension Systems and Reform") in 2005. Two new pillars were added in the form of Pillar 0 and Pillar 4. Pillar 0 was non-contributory and was meant for poverty reduction and was meant to be means tested. Pillar 4 reflected the importance of the informal sector in developing and low-income countries. The five pillars as proposed by the World Bank are summarised in Table 8.

	Target Group			*Main Criteria*		
Pillar	*Lifetime Poor*	*Informal Sector*	*Formal Sector*	*Characteristics*	*Participation*	*Funding or Collateral*
0	**X**	X	x	"Basic" or "social pension," at least social assistance (universal or means tested)	Universal or Mandated	Budget or general revenues
1			**X**	Public pension plan, publicly managed (defined benefit or notional defined contribution)	Mandated	Contribution, perhaps with some financial reserves
2			**X**	Occupational or personal pension plans (fully funded defined benefit or fully funded defined contribution)	Mandated	Financial assets
3	x	**X**	**X**	Occupational or personal pension plans (partially or fully funded defined benefit or funded defined contribution)	Voluntary	Financial assets
4	**X**	**X**	X	Access to informal support (family), other formal social programs (health care), and other individual financial and nonfinancial assets (homeownership)	Voluntary	Financial and nonfinancial Assets

Note: The size and appearance of x reflect the importance of each pillar for each target group in the following increasing order of importance: x, X, **X**.

Table 8: Multipillar Pension Taxonomy (Source: WB, 2005)

Whilst the reform of pensions systems and its implications for economic aggregates have attracted a lot of empirical attention since the publication of "Averting the Old Age Crisis", little has been said on the redistributive issues both from a theoretical and empirical perspective. Our key contribution thus comes from the perspective of the welfare issues in the design of pensions systems. Unsurprisingly, the pensions reforms in Latin America has attracted a lot of attention. Holzmann (1997) considers the impact of pension reforms on financial market development and economic growth in Chile; Bertranou et al. (2004) look at the impact on poverty reduction in 5 Latin

American countries; Corbo and Schmidt-Hebbel (2003) consider the macroeconomic effects of pensions reforms in Chile; Edwards and Cox-Edwards (2002) look at the impact on labour markets in Chile; Packard (2002) and Gill et al. (2004) consider the Latin American experience with pensions reforms. Yermo (2002) considers the performance of funded pensions systems in Latin America. Bailliu and Reisen (1997), Bosworth and Burtless (2003) and Antolin et al. (2004) look at pensions reforms and savings; James et al. (2003) investigates the gender issues in pensions reforms; Vittas (2000) and Walker and Lefort (2001) consider reforms and capital market development.

In this paper, we aim to formalise the intuition behind the WB model and seek to identify the mechanisms through which the pillars affect the macroeconomy and welfare. We also consider some of the redistributive issues involved with the design of pensions systems and the implications for welfare. Using a two period overlapping generations general equilibrium model, we design a three-pillar pensions system. The aim is very much as proposed by the World Bank except for one difference. The model consists of heterogeneous agents who differ according to their productivity (high/low) and potentially their discount factors. Each agent pays a proportion of wages in tax whilst working and expects to receive a pensions in retirement. The pensions the agent receives depends on Pillars 1, 2 and 3. Under Pillar 1 all agents receive the same level of pensions, irrespective of their type and level of contribution. Under Pillar 2 the pensions an agent receives is proportional to his contribution The agent then complements the pensions received with his own voluntary savings - Pillar 3. Since Pillar 1 is redistributive, the weight the planner attaches to Pillar 1 depends on his inequality-aversion. The higher the aversion, the more the weight attached to Pillar 1. The low productivity agent (the poor) prefers more redistribution than the high productivity agent (the rich); he will thus always prefer a higher weight on Pillar 1. On the other hand, the rich agent will prefer Pillar 2. Once the planner has chosen the weight, it finds the optimal level of tax rate consistent with the competitive equilibrium.

The only difference with the WB approach is that in this model, the second Pillar is publicly managed. Those who argue in favour of Pillar 2 being privately managed do so on the premise that returns from the privately managed portfolio (Pillar 2) tends to be higher than the publicly managed portfolio (Pillar 1). However, Barr (2001) argues that there is no reason for the privately managed portfolio to yield a higher return once the risks and administration costs are included. In this respect, our assumption that Pillar

1 and Pillar 2 have the same rate of return is not unreasonable and does not affect our general results in a significant manner. The model is deliberately kept simple to allow for analytical tractability in so far as possible.

The key findings of this chapter can be summarised as follows: The type of pensions system in place and the weight attached to each pillar can have a significant impact for capital formation and the welfare of the agents. Some of the key results that we are able to derive analytically suggest:

1. An increase in the weight attached to Pillar 1 increases the welfare of the poor (UP) but reduces the welfare of the rich (UR). It also hinders capital accumulation.

2. An increase in the proportion of the rich, and their productivity, leads to an unambiguous increase in the welfare of the rich, the poor and aggregate welfare.

3. An increase in the productivity of the poor leads to an unambiguous increase in the welfare of the rich, the poor and aggregate welfare. The welfare of the poor increases by more than that of the rich.

4. A decrease in the population growth rate (ageing) leads to an unambiguous marginal fall in the welfare of the rich, the poor and aggregate

welfare.

To derive the optimal tax rate, we take recourse to simulations. The key finding suggests that, for some plausible range of parameters, whether the planner decides to attach a higher weight to Pillar 1 or to Pillar 2 makes a marginal difference to the optimal tax rate. The other findings can be summarised as:

1. An increase in the weight attached to Pillar 1 leads to a marginally higher tax rate and a lower steady state capital. However, since the welfare of the poor rises by more than the fall in the welfare of the rich, aggregate welfare increases.

2. As the proportion of rich increases, the tax rate increases (suggesting richer economies can afford higher pensions). Steady state capital and the welfare of the rich increase as well.

3. The productivity of the poor is non-linear with respect to the tax rate. When the productivity of the poor is very low, the tax rate is very high. However, as the productivity increases, this causes the tax rate to fall and as the productivity increases further, the tax rate rises gradually.

As the productivity of the poor increases, capital and aggregate welfare increase as well.

The economic intuition is as follows: An increase in the tax rate leads to a fall in steady-state capital but leads to an increase in the utility of the poor. At low levels of productivity, the increase in the utility of the poor is greater than the cost in terms of capital formation. Beyond a certain level, the cost in terms of capital formation exceeds the gains in the utility of the poor and the tax rate falls. At high levels of productivity, the increase in the tax rate ensures the economy remains dynamically efficient.

4. The impact of population ageing on the tax rate is infinitesimal. Though the tax rate changes infinitesimally as the population ages and capital increases, aggregate welfare falls.

5. If both sets of agents are equally impatient, the tax rate and capital falls whilst aggregate welfare increases. On the other hand, if only the poor are impatient, there is a non-linear impact on capital. The tax rate falls and beyond a certain point it is no longer optimal to have a social security system. Though aggregate welfare increases, this gives

rise to a situation where the welfare of the poor can exceed that of the rich.

The rest of this chapter is as follows: in Section 2 we describe the set up of the economy. Section 3 derives the competitive equilibrium whilst Section 4 considers the planner's problem in terms of finding the optimal tax rate. We then test the model to data from 23 countries. Section 6 concludes and provides some potential extensions.

5.2 The Economy

We consider an infinitely lived overlapping generations economy in the Samuelson-Diamond spirit. Economic activity takes place over discrete time $t \in \{0, 1..., \infty\}$ and there is no uncertainty. The economy consists of two types of utility maximising agents, h, heterogeneous in their productivity ψ and possibly their discounting of the future; profit maximising perfectly competitive firms; and a welfare maximising social planner. The planner maximises the welfare of all agents born at time t. At each time t, two factors, an amount of capital, k, and labour, l, are available as inputs to production and a homogeneous good, y, is produced.

5.2.1 Agents

At time t, one young generation and an old generation live simultaneously. Population grows at a constant rate n. Hence, at any time t, there are $(1+n)$ more (young) workers than (old) retirees. The population at t can thus be expressed as: $L_t + L_{t-1} = (2+n)L_t$. L_t refers to the agents born at t. Following Samuelson (1958), the distribution of the population is considered to be stationary. Therefore, the proportions and types of individuals remain the same across generations.

When young, each agent provides one unit of labour inelastically. When old the labour endowment is zero. There are two types of agents who differ in terms of their productivity $\psi \in (0,1]$ and possibly their discounting of the future. For some cases, we assume in line with Becker (1990), that the poor can potentially attach a lower weight to their future consumption than the rich. In competitive markets, the productivity determines the wages the agents receive. We classify the high productivity agents as rich and the low productivity agents as poor. The rich and the poor make up proportions π and $(1-\pi)$ of the economy respectively. For ease of manipulation, let us normalise the productivity of the rich to 1 such that any agent with $\psi < 1$ as poor.

Agents maximise utility by maximising consumption $u^h(c_t^h, c_{t+1}^h)$ subject to the budget constraints over their lifetime. They are non-altruistic: they are born without any assets and leave no bequests. When young, the agents choose the levels of consumption and savings which maximise their utility, whilst when old the agents live off their savings and pensions. For ease of manipulation, the intertemporal utility function is taken to be additive and log-linear. The utility function is thus strictly concave, since more consumption is preferred to less, and twice differentiable: $u'(c) > 0$ and $u''(c) < 0$. The function also satisfies $\lim_{c \to 0} u'(c) = \infty$ such that subject to its disposable income, the household will always choose a positive level of consumption when maximising life-cycle utility. We can express the utility-maximising problem of a rich agent as follows:

$$\begin{aligned} \underset{\{c_t^y, c_{t+1}^o, s_t\}}{Max} &: \quad u^h = \ln c_t^{y,h} + \beta^h \ln c_{t+1}^{o,h} && (1) \\ \text{subject to} &: \\ c_t^{y,h} &= \psi^h w_t (1 - \tau) - s_t^h && (2) \\ c_{t+1}^{o,h} &= R_{t+1} s_t^h + P_{t+1}^h && (3) \end{aligned}$$

Eqn. (1) suggests that the agents maximise utility over both periods of their lifetime and future consumption is discounted by a factor $\beta^h \in (0,1)$[11]. The budget constraints suggest when young, the agent earns a productivity-related wage $\psi^h w_t$ out of which a constant proportion τ is taxed. In line with Diamond (1965), the disposable income is then either consumed or saved. When old, the agent consumes his savings and the interest rates earned thereon plus the pensions P^h_{t+1} he receives. P^h_{t+1} is based on the weight the planner allocates to Pillars 1 and Pillars 2. We discuss this further in the next section.

5.2.2 Firms

A large number of identical firms produce a homogeneous good using an identical economy-wide Cobb Douglas production function. The production function can be represented as $Y = K^\alpha L^{1-\alpha}$, where α is the share of capital in production.

In intensive form, this production function reduces to $y = k^\alpha$. Firms maximise profit by taking factor prices, which are paid their marginal products in a competitive setting, as given. It is assumed that the labour market clears such that labour demand equals labour supply and the wages received by a worker depends on his level of productivity. The economy is endowed with an initial capital stock $K_0 > 0$ and capital depreciates fully from one period to the next. We assume no technological change. The production function satisfies the usual conditions such that $f(0) = 0$, $f'(k) > 0$, $f''(k) < 0$ and the Inada conditions: $\lim_{k \to 0} f'(k) = \infty$ and $\lim_{k \to \infty} f'(k) = 0$.

5.2.3 Social Planner

The social planner maximises the welfare of all agents born at time t. The planner chooses to impose a tax on the wages of the young agents working and allocate the proceeds to those who are retired at time t. The tax is imposed on both the rich and the poor at a rate τ. This rate is assumed to be constant through time. Hence, the total tax revenue raised by the planner is $T = \pi L_t \tau w_t + (1 - \pi) L_t \tau \psi w_t$ and this can be expressed as:

$$T_t = [\pi + (1 - \pi) \psi] \tau L_t w_t \tag{4}$$

Under a balanced budget rule, the planner determines which proportions $\omega \in (0, 1)$ to allocate to Pillar 1 and the remainder $(1 - \omega)$ to Pillar 2. ω can be seen as a reflection of the degree of inequality aversion of the planner. The higher the value of ω, the higher the perceived inequality aversion, and the higher the degree of intragenerational redistribution.

Pillar 1 Pillar 1 ensures that everyone has a minimum level of income in retirement and it is allocated to all the old agents irrespective of their level of productivity. Hence, the total amount allocated to Pillar 1 is P^1 and this is allocated in proportions π and $(1-\pi)$ to the first pillar pensions of the rich $(P^{1,r} = L_{t-1}p_t^1)$ and the poor $(P^{1,p} = L_{t-1}p_t^1)$, respectively:

$$P^1 = \omega T = \pi P^{1,r} + (1-\pi) P^{1,p} \tag{5}$$

Hence, each agent is guaranteed to receive p^1 in retirement under Pillar 1. Since there is no distinction according to ability, each agent receives the following in pensions under the first pillar:

$$p_t^1 = \omega (1+n) [\pi + (1-\pi) \psi] \tau w_t \tag{6}$$

Pillar 1 has a redistributive component in that although all the agents pay the taxes on their wages (which is based on their ability), the contribution is not taken into consideration under Pillar 1. The poor are relatively better off under this scheme. There is thus an element of intragenerational redistribution that takes place under Pillar 1 since the rich agents, who contribute more that the poor agents end up receiving the same level of pensions as the poor.

Pillar 2 Pillar 2 provides an additional pension to an agent proportional to his contribution. The contribution rates, γ, of the rich and the poor can be expressed as a function of the taxes they paid compared to the total tax (T) raised. γ can be expressed as:

$$\gamma^r = \frac{\pi L_t \tau w_t}{[\pi + (1-\pi)\psi] L_t \tau w_t} = \frac{\pi}{[\pi + (1-\pi)\psi]} \tag{7}$$

$$\gamma^p = \frac{(1-\pi)\psi L_t \tau w_t}{[\pi + (1-\pi)\psi] L_t \tau w_t} = \frac{(1-\pi)\psi}{[\pi + (1-\pi)\psi]} \tag{8}$$

The total payments under pillar 2 is $(1-\omega)T$ and this is allocated to the rich and the poor in the proportions of γ^r and γ^p, respectively. Hence, total pensions paid under Pillar 2, P^2, and the amount received by the old rich, $p^{2,r}$, and the poor, $p^{2,p}$, can be expressed as:

$$P^2 = (1-\omega)T = \frac{\pi(1-\omega)T}{[\pi + (1-\pi)\psi]} + \frac{(1-\pi)\psi(1-\omega)T}{[\pi + (1-\pi)\psi]} \tag{9}$$

$$p^{2,r} = \frac{\gamma^r(1-\omega)T}{\pi L_{t-1}} = (1+n)(1-\omega)\tau w_t \tag{10}$$

$$p^{2,p} = \frac{\gamma^p(1-\omega)T}{(1-\pi)L_{t-1}} = (1+n)(1-\omega)\tau\psi w_t \tag{11}$$

Under Pillar 2, we note that the pensions the agents receive is a function of their contribution. Comparing Eqns (10) and (11) above, it can be seen that the pensions the agents receive under Pillar 2 is a function of their abilities. There is no intragenerational redistribution under this pillar.

Pillar 3 Pillar 3 is entirely voluntary and seeks to ensure that the agents save a proportion of their income voluntarily. Though this could be encouraged by allowing for tax deductions or the use of other instruments, we allow this pillar to be entirely voluntary and dependent on the utility maximisation decision of the agents. In this framework, government policy does not provide explicit incentives for the promotion of Pillar 3. However, we cannot rule out the fact that government policy to promote Pillars 1 and 2 can lead to a crowding out of Pillar 3.

Total Pensions We restrict the definition of total pensions as that accruing under Pillars 1 and 2. Pillar 3 is accounted for separately. The total pensions received by the agents consists of the payments under Pillars 1 and 2. Hence, $P^h = p^1 + p^{2,h}$. The total pensions received by the rich and the poor are:

$$P^r_{t+1} = [\omega (1-\pi)(\psi - 1) + 1](1+n)\tau w_{t+1} \tag{12}$$

$$P^p_{t+1} = [\omega\pi (1-\psi) + \psi](1+n)\tau w_{t+1} \tag{13}$$

In the above equations, ω determines the extent to which the pensions system is redistributive. If $\omega = 1$, there is no Pillar 2 and the system is completely redistributive. On the other hand, if $\omega = 0$, pensions are simply based on ability - and there is no Pillar 1.

5.3 Competitive Equilibrium

Given the households' and the firms' objectives, a competitive equilibrium for the economy can be defined as a sequence of consumption $\{c_t^y, c_t^o\}_{t=0}^{\infty}$ such that:

1. Given a sequence of taxes and transfers, $\{\tau\psi^h w_t, P_t^h\}_{t=0}^{\infty}$, and the prevailing competitive wages, w_t, and interest rate, R_t, solve the individual's optimisation problem subject to satisfying the Euler equation;

2. Factors of production are paid their marginal products ($w_t = (1-\alpha)k_t^{\alpha}; R_t = \alpha k_t^{\alpha-1}$) and labour and capital markets clear such that $L_t^D = L_t$ and $S_t = K_{t+1}$;

3. The planner's budget is always balanced hence taxes raised is redistributed between Pillars 1 and 2: $T = P^1 + P^2$;

4. The economy's resource constraint is always satisfied. In intensive form,

the constraint is defined as the allocation of current output, y_t, between consumption for the two generations living simultaneously and savings.

$$y_t = c_t^y + \frac{c_t^o}{(1+n)} + (1+n)k_{t+1} \tag{14}$$

Given the above definition of competitive equilibrium, we can express the intertemporal budget constraints for the rich and the poor as:

$$c_t^{y,h} + \frac{c_{t+1}^{o,h}}{R_{t+1}} = \psi^h w_t(1-\tau) + \frac{P_{t+1}^h}{R_{t+1}} \tag{15}$$

We can then write the Lagrangian as:

$$\underset{\{c_t^y, c_{t+1}^o\}}{Max} : \ell = \ln c_t^{y,h} + \beta^h \ln c_{t+1}^{o,h} - \lambda[c_{t+1}^{o,h} - R_{t+1}\left\{\psi^h w_t(1-\tau) - c_t^{y,h}\right\} + P_{t+1}^h] \tag{16}$$

and derive the standard Euler equation which dictates the optimal allocation of consumption over the two periods of an agents lifetime. β^h refers to the discount factor of the agents. The Euler equation is:

$$c_{t+1}^o = \beta^h R_{t+1} c_t^y \tag{17}$$

We can now proceed to derive the optimal levels of consumption and savings of both sets of agents. The optimal levels of consumption and savings of both sets of agents are shown in Appendix 5.1.

The optimal levels of consumption and savings of the agents depend on their disposable income and the discounted value of the pensions received. We find that there is potential crowding out that arises from the provision of pensions under Pillar 1 which crowds out (private) savings - which determines Pillar 3. We also note that the higher the value of ω, the greater the degree of crowding out. As such, there is greater crowding out under Pillar 1 - which is more redistributive - that under Pillar 2. The optimal level of savings of the agents can be seen in Eqns $(18) - (19)$

$$s^{y,r} = \frac{1}{[1+\beta]}\left[\beta w_t(1-\tau) - \frac{[\omega(1-\pi)(\psi-1)+1](1+n)\tau w_{t+1}}{R_{t+1}}\right] \quad (18)$$

$$s^{y,p} = \frac{1}{[1+\theta]}\left[\theta\psi w_t(1-\tau) - \frac{[\omega\pi(1-\psi)+\psi](1+n)\tau w_{t+1}}{R_{t+1}}\right] \quad (19)$$

Based on Eqns $(18) - (19)$, we can proceed to derive the steady-state capital stock in the economy. We assume complete depreciation[12] of capital from one period to the next. As such, capital formation is a function of Pillar 3. Hence, although there is crowding out , the aggregate level of savings in the economy should be positive. We can then characterise capital formation

[12] In the non-conventional sense: the rich consume capital and the young replenish it.

as the savings of the young rich and poor, i.e., $S_t = S_t^r + S_t^p = K_{t+1}$. In intensive form, this reduces to:

$$(1+n)\,k_{t+1} = \pi s_t^r + (1-\pi)\,s_t^p \tag{20}$$

Hence, in steady state where $k_{t+1} = k_t = k^*$, the capital available per unit of worker is a function of $(\beta, \theta, \pi, \psi, \alpha, n, \tau, \omega)$. The steady state capital $k^*(\beta, \theta, \pi, \psi, \alpha, n, \tau, \omega)$ is thus affected not only by the pensions system, but also by proportions of rich and poor in the economy; the level of productivity, and the usual parameters pertaining to the discount factor, share of capital in production; and the population growth rate. When the agents have the same discount factor such that $\beta = \theta = \delta$, the capital per unit of worker is:

$$k^* = \left[\frac{\delta\left\{\pi + (1-\pi)\,\psi\right\}\alpha(1-\alpha)(1-\tau)}{(1+n)\left\{(1+\delta)\,\alpha + \left[\pi + (1-\pi)\,\psi\right](1-\alpha)\tau\right\}}\right]^{\frac{1}{1-\alpha}} \tag{21}$$

In this case, $k^*(\delta, \pi, \psi, \alpha, n, \tau)$ and ω does not matter. Hence, when the agents have the same discount factor, the weight attached to Pillars 1 and 2 does not matter. On the other hand, we can see from eqn. (1) that since ω appears only in the denominator, a fall in ω (that is an increased emphasis on Pillar 2) leads to an increased level of k. We consider the Eqn. (21) where agents have the same discount factor to do some comparative statics on k.

Proposition 12 ***(a) An increase in*** δ ***leads to an increase in*** k

(b) An increase in π ***leads to an increase in*** k

(c) An increase in ψ ***leads to an increase in*** k

(d) An increase in τ ***reduces*** k

(e) An increase in n ***reduces*** k

Proof. The Proof is in Appendix 5.1. ■

We consider the general steady state capital Eqn. 1 (Appendix 5.1) to consider the impact of ω on k.

Proposition 13 ***An increase in*** ω ***leads to a fall in*** k ***if*** $\beta > \theta$***. There is no impact if*** $\beta = \theta$***.***

Proof. The Proof is in Appendix 5.1. ■

Since we are only concerned with the case of $\beta > \theta$ and $\beta = \theta$, an increase in ω can only lead to a reduction in k or have no impact on k, depending on β and θ. This confirms our hypothesis that the higher the weight attached to Pillar 1, the greater the extent of crowding out.

5.4 The Planner's Problem

For a given aversion to inequality, ω, the planner maximises social welfare by maximising the utility of all agents born at time t. As discussed earlier, the aversion to inequality is determined by the weight attached to Pillars 1. The higher the inequality aversion of the planner, the greater the extent of redistribution undertaken and the higher the weight attached to Pillar 1. To achieve its objective for a given ω, the planner considers the competitive equilibrium and chooses τ to maximise $\{c_t^y, c_{t+1}^o, k_{t+1}\}_{t=0}^{\infty}$ subject to the allocation $\{y_t\}_{t=0}^{\infty}$. In steady state, the planner's problem can be expressed as:

$$\underset{\{\tau\}}{Max} : V = \sum_{t=0}^{\infty} \rho^t \left(U\left(c^r\right) + U\left(c^p\right) \right) \tag{22}$$

We write the planner's problem as that of maximising the utility of all agents born at time t. $\rho \leqslant 1$ indicates the weight the planner attaches to future generations. We can then write the welfare function in intensive form as:

$$V = \pi \left[\ln(c^{y,r}) + \beta \ln(c^{o,r})\right] + (1 - \pi) \left[\ln(c^{y,p}) + \theta \ln(c^{o,p})\right] \tag{23}$$

The first part of eqn. (23) refers to the proportion of rich born at time any point in time in the economy. The second part considers the poor and their weight in the economy. The welfare function can then be expressed as a function of the lifetime consumption of all the agents. The consumption levels follow from Appendix 5.1.

We can then express V as a function of the policy-parameters ω and τ and non-policy parameters $(\alpha, \beta, \theta, \pi, \psi, n)$

$$\begin{aligned} V &= U \ln\left[J + [L + M\omega]\,\tau\right] + Z \ln\left[N + [O + P\omega]\,\tau\right] \\ &\quad + Q\left[\ln A(1-\tau) - \ln\left[B + \{G + H\omega\}\,\tau\right]\right] \end{aligned} \tag{24}$$

where A, B, G, H, J, Q, U, L, M and Z are all non-policy parameters and defined in Appendix 5.1.

The impact of π, ψ and n on Welfare Given the general welfare function (24), we can consider the impact of π, ψ and n on aggregate welfare. For ease of manipulation, we assume that $\tau = \omega = 0$ and the agents have the same discount factor $\beta = \theta = \delta$. We can then rewrite the aggregate welfare function, V, and the utility functions of the rich, UR, and poor, UP, as:

$$
\begin{aligned}
V &= z\left[\ln\left[\pi + (1-\pi)\,\psi\right] + (1+\delta)\,(1-\pi)\ln\psi - \ln(1+n)\right] + z_1{}^{13} \\
UR &= z\left[\ln\left[\pi + (1-\pi)\,\psi\right] - \ln(1+n)\right] + z_2{}^{14} \\
UP &= \left[\ln\left[\pi + (1-\pi)\,\psi\right] - \ln(1+n)\right] + (1+\delta)\ln\psi + z_3{}^{15}
\end{aligned}
$$

where $z = \left[\frac{(1+\delta)\alpha+\delta(\alpha-1)}{(1-\alpha)}\right]$ and z_1, z_2 and z_3 refer to parameters independent of π, ψ and n.

Proposition 14 ***(a) An increase in*** π ***increases welfare unambiguously***

(b) For δ, ***an increases in*** π ***increases welfare of the rich and the poor by the same extent***

Proof. The Proof is in Appendix 5.1. ■

Although in the case where the agents have the same discount factor, the increase in the welfare of the rich and the poor is the same when the proportion of rich increases, this might differ if the agents do not have the same discount factor.

Proposition 15 ***(a) An increase in*** ψ ***increases aggregate welfare unambiguously***

(b) For an increase in ψ, the welfare of the poor increases by more than that of the rich.

Proof. The Proof is in Appendix 5.1. ■

Proposition 16 ***(a) An increase in n decreases welfare unambiguously***

(b) For δ, an increase in n reduces the welfare of the rich and the poor by the same extent

Proof. The Proof is in Appendix 5.1. ■

The impact of ω and τ on Welfare of Rich and Poor From the aggregate welfare function, still assuming the same discount factor, we can express the utility of the rich and that of the poor as:

$$UR = \begin{bmatrix} \left[\frac{(1+\delta)\alpha-(\alpha-1)}{(1-\alpha)}\right][\ln(1-\tau)-\ln(b+c\tau)] \\ +(1+\delta)\ln\left[\alpha\left[b+c\tau\right]+a\left[\omega\left(1-\pi\right)\left(\psi-1\right)+1\right]\left(1+n\right)\tau\right]+{z_3}^{16} \end{bmatrix}$$

$$UP = \begin{bmatrix} \left[\frac{(1+\delta)\alpha-(\alpha-1)}{(1-\alpha)}\right][\ln(1-\tau)-\ln(b+c\tau)] \\ +(1+\delta)\ln\left[\alpha\psi\left[b+c\tau\right]+a\left[\omega\pi\left(1-\psi\right)+\psi\right]\left(1+n\right)\tau\right]+{z_4}^{17} \end{bmatrix}$$

where:

$a = \delta \{\pi + (1-\pi)\psi\} \alpha(1-\alpha);$

$b = (1+\delta)(1+n)\alpha;$

$c = [\pi + (1-\pi)\psi](1+n)(1-\alpha).$

Proposition 17 ***The utility of the rich (poor) is decreasing (increasing) in*** ω

Proposition 18 ***The impact of*** τ ***depends on whether*** $\frac{\partial UP}{\partial \tau} \gtreqless \frac{\partial UR}{\partial \tau}$.

Proof. The Proof is in Appendix 5.1. ■

If $\frac{\partial UP}{\partial \tau} > \frac{\partial UR}{\partial \tau}$, this implies that the increase on the utility of the poor is higher than that of the rich. Hence, the tax improves aggregate welfare.

5.4.1 The Optimal Level of τ

To find the optimal tax rate consistent with ω and the competitive equilibrium, we proceed to differentiate V with respect to τ in Eqn. (24). This yields:

$$\frac{\partial V}{\partial \tau} = \frac{U[L+M\omega]}{[J+[L+M\omega]\tau]} + \frac{Z[O+P\omega]}{[N+[O+P\omega]\tau]} - \frac{Q}{[1-\tau]} - \frac{Q[G+H\omega]}{[B+\{G+H\omega\}\tau]} = 0 \tag{25}$$

where the non-policy parameters are defined as before. Given the nature of the Eqn. (25), we have recourse to simulations to derive the optimal level of τ.

5.4.2 Simulation Results and Sensitivity Analysis

We now consider the behaviour of the model for a given set of parameters. These parameters have been chosen to initially yield a dynamically efficient solution and provide plausible values of τ for $\omega \in (0, 1)$. We assume the agents differ in terms of their productivity and discount factor, which is taken to be line in line with values used for macro simulations (see de la Croix and Michel, 2002; Krueger and Kubler, 2006). We initially set the discount factor at 0.96 for both agents, though when carrying out the sensitivity analysis we set the discount factor of the poor to 0.9. The population growth rate is set at 1 percent. In the baseline, we assume that the poor earn 75 percent of what the rich earn and we assume that the rich, π, make up 40 percent of the economy, with the poor accounting for the remaining 60 percent. We carry out sensitivity analysis on all the parameters.

Baseline Results The baseline results are presented in Table 9:

		Alpha	Beta	Theta	Pi	Psi	n	
		0.25	0.96	0.96	0.4	0.75	0.01	
Omega	Capital	Output	Wages	Interest	Tau	UR	UP	V
0	0.14018	0.61189	0.45892	1.0912	0.13184	-2.8203	-3.3841	-3.1586
0.5	0.13944	0.61108	0.45831	1.0956	0.1336	-2.8387	-3.368	-3.1562
1	0.13881	0.61039	0.45779	1.0993	0.13511	-2.8576	-3.3517	-3.154
		Alpha	Beta	Theta	Pi	Psi	n	
		0.25	0.96	0.9	0.4	0.75	0.01	
Omega	Capital	Output	Wages	Interest	Tau	UR	UP	V
0	0.14044	0.61217	0.45913	1.0897	0.12244	-2.819	-3.2797	-3.0954
0.5	0.13975	0.61142	0.45857	1.0937	0.124	-2.836	-3.2652	-3.0935
1	0.13916	0.61078	0.45808	1.0972	0.12532	-2.8535	-3.2505	-3.0917

Table 9: Baseline Simulations

The first part of the results in Table 9 consider the situation where the agents have the same discount factor $\beta = \theta$ whilst the second part considers the case where $\beta > \theta$. When $\omega = 0$, there is only Pillar 2 and the pension is totally ability based. On the other hand, if $\omega = 1$ the pension is under Pillar 1 and is completely redistributive. For $\omega = 0.5$, an equal weight is attached to Pillars 1 and 2. In both cases, we find that as ω tends to 1, steady state capital falls. Since the weight attached to each pillar affects the capital stock, it also has implications for output, wages and interest rates. Hence, an economy with a higher weight to Pillar 1 will tend to have a higher interest rate than an economy with a higher weight on Pillar 2. Thus, the more redistributive the system, the higher the interest rate. We can thus interpret redistribution as a distortion in a dynamically efficient economy which hinders capital formation.

We also find that the tax rate for the current set of parameters is around 12-13 percent. The tax rate is higher when $\beta = \theta$ than when $\beta > \theta$, and this is in line with previous findings. The intuition behind this result follows from the fact that steady-state capital is greater for $\beta = \theta$ than $\beta > \theta$ thereby suggesting a richer economy can afford more generous pensions.

The results in so far as the welfare of the rich and the poor are concerned suggests that the welfare of the poor is higher the higher the value of ω, ie, the higher the redistributive component of the pensions system. On the other hand, the welfare of the rich falls as ω increases. These results are consistent with the earlier proposition suggesting $\frac{\partial UR}{\partial \omega} < 0$ and $\frac{\partial UP}{\partial \omega} > 0$. Aggregate welfare is higher as ω increases and $\beta > \theta$. The overall impact of ω on welfare can be summarised as follows:

- An increase in ω leads to an increase in the tax rate;
- The increase in ω and the tax rate reduces the capital stock (with implications for output, wages and the interest rate);

- The increase in the tax rate reduces the welfare of the rich, whilst increasing the welfare of the poor (through the redistributive component);

- The overall impact depends on whether the welfare loss arising from the fall in capital and the welfare of the rich is compensated for by the gain in the welfare of the poor.

Sensitivity Analysis The baseline results gave an overview of the optimal tax rate and implications for welfare for given values of ω and a set of parameters. However, to get a better understanding of the behaviour of the endogenous variables, we now allow for the exogenous parameters to vary. This is done with the aim of getting a better understanding of how the policy variables behave with respect to changes in the selected parameters. We change one parameter at a time and keep all the others constant. We note that for some of the parameters, it is no longer optimal to have a pensions policy beyond a certain point.

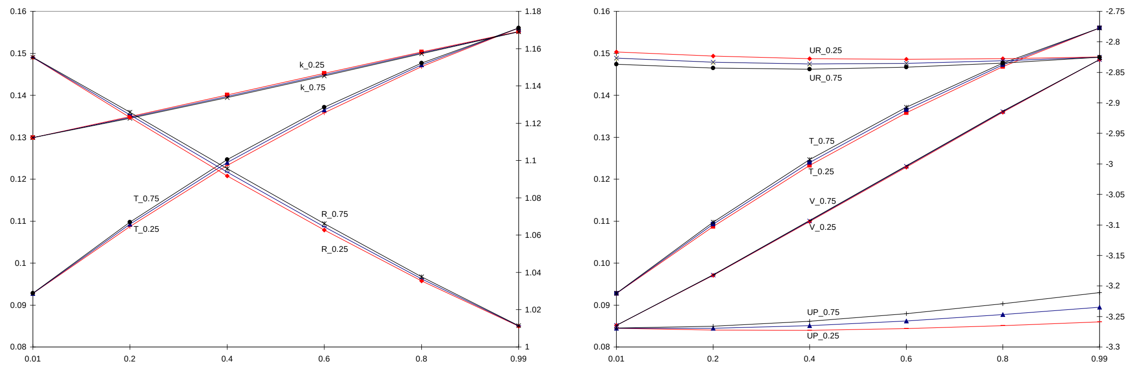


Figure 14: Impact of changes in Pi

Proportion of Rich When there is an increase in the proportion of rich, capital increases unambiguously, however, in line with previous propositions, we notice that for higher levels of ω, the capital stock is lower. Hence, as the proportion of rich increases, the increase in capital leads to higher output and wages as well as lower interest rates. However, the impact on (capital and) interest rate is much lower than before in that for instance when the proportion of rich increases from 1% to 20%, the interest rate falls from around 15.5% to around 12.3-12.6%. The economy remains dynamically efficient throughout.

In so far as the tax rate is concerned, as the proportion of rich increases, the tax rate increases simultaneously. This suggests that richer economies are able to afford more generous pensions. For instance, when the proportion of rich is 1%, the tax rate is around 9.2%. However, as the proportion of rich rises, the tax rate rises as well, to around 15.6% when the proportion of rich rises to 99%. The increase in the tax rate can be seen as a mechanism that ensures the economy remains dynamically efficient at all times even when the proportion of rich tends to 1. As before, the tax rate remains marginally higher the higher the redistributive component.

Whilst as the proportion of rich increases welfare increases unambiguously, the welfare of both the rich and the poor initially falls. However, beyond a certain point, the welfare of both groups increases as the proportion of rich goes up. This turning point is reached faster for the poor than the rich. We can try and explain some of this mechanism as follows:

- As π increases, the tax rate goes up but so does the capital stock;
- Initially, the increases in the tax rates exceeds the gains from the increase in capital;
- However, for the poor, the increase in tax rates increases aggregate welfare faster through the redistributive component;
- Eventually, for both groups, the increase in capital eventually exceeds the rise in taxes.

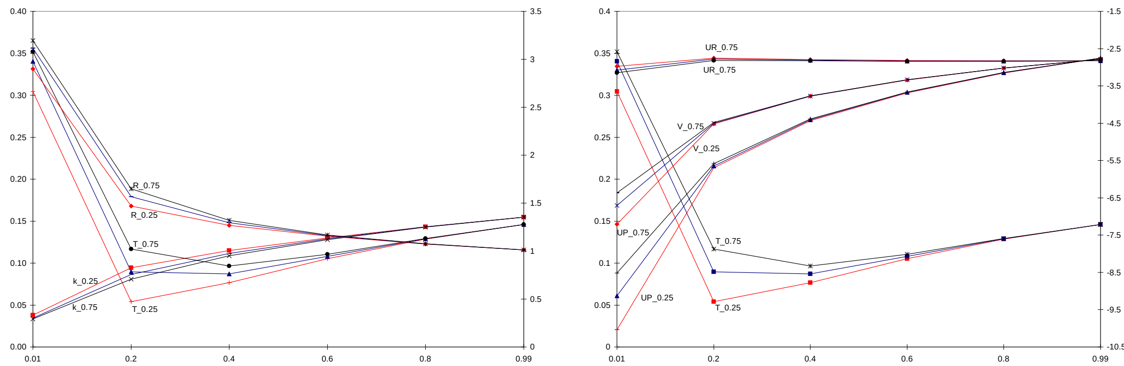


Figure 15: Impact of changes in Psi

Productivity of Poor

The productivity of the poor yields some interesting implications in so far as the key endogenous variables are concerned. In line with our earlier proposition, as the productivity of the poor increases, the capital stock increases unambiguously. The productivity of the poor has a significant impact on the interest rate but the economy remains dynamically efficient throughout. As before, we note that capital is higher when there is a lower weight on the first Pillar. This is in line with our previous discussion suggesting a high level of redistribution hinders capital accumulation.

Compared to capital, the behaviour of the tax rate is somewhat unexpected. When the productivity of the poor is very low the tax rate is fairly high (example when the productivity of the poor is 1% that of the rich, the tax rate varies between a high of 35% for $\omega = 0.75$ and 30% for $\omega = 0.25$). However, as the productivity of the poor increases to 20% that of the rich, the tax rate falls in all cases varying between 5.4% and 11.6%. The tax rate then turns up (at different points) as the productivity of the poor increases further. When the productivity of the poor is 99% that of the rich, the tax rate is around 14.6%. This lends further credence to our suggestion that richer economies are able to afford more generous pensions. As before, the tax rate is higher the higher the weight attached to Pillar 1.

As the productivity of the poor increases, aggregate welfare increases unambiguously and the same applies to the utility of the poor. However whilst the utility of the rich initially improves due to the falling taxes, it then falls for a while (as the tax starts to rise) before picking up when the productivity of the poor is fairly high. The increase in the welfare of the rich suggest the gains from the increases in capital exceed the cost of the rising taxes. Unlike the poor, we note that the welfare of the rich is highest when a higher weight is attached to Pillar 2 (could also be due to the fact that this yields a lower tax rate). The mechanism could be as follows:

- As ψ increases, the capital stock rises gradually whilst the tax rate initially falls and this increases the welfare of the rich;

- However, the tax rate then starts to rise gradually and this exceeds the gain from the rise in capital;

- Eventually, as the capital stock continues to rise, this exceeds the cost of higher taxes and the welfare of the rich picks up again.

Population Growth Rate One of the main reasons the World Bank suggested the introduction of a three-pillar based pensions system was to mitigate some of the adverse effects of an ageing population on the fiscal position of governments. Though this model considers a balanced budget rule, we can also consider the implications of an ageing population on taxation and aggregate welfare. With an ageing population, there is a marginal increase in capital available per worker and this in turn results in a decrease in the interest rate. Output and wages increase marginally. However, the results pertaining to the tax rate yield some interesting insights. The ageing population does not have any (major) impact on the tax rate. In fact as the population ages, the tax rate falls infinitesimally. The earlier results in so far as welfare is concerned continue to hold in that the higher the redistributive component, the higher the welfare of the poor whilst the welfare of the rich

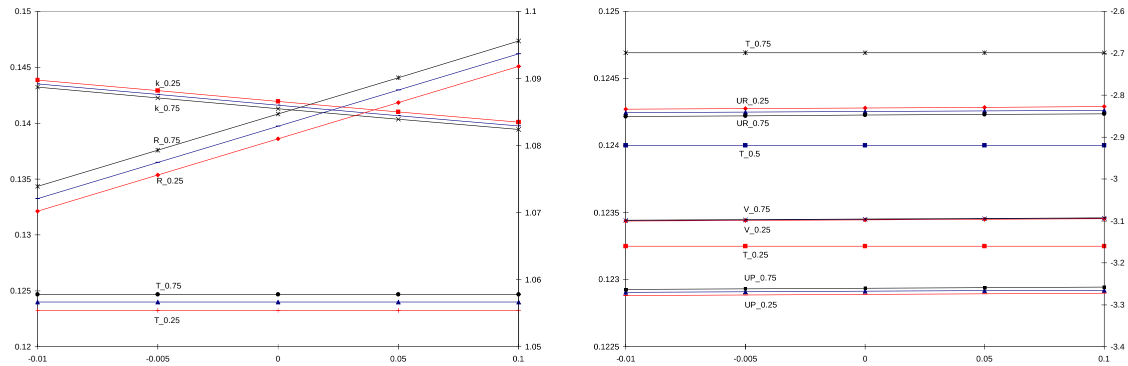


Figure 16: Impact of Change in Population Growth Rate

is lower. However, as the population ages, the welfare of both the rich and the poor falls infinitesimally. Aggregate welfare is also infinitesimally lower. Some of the results can be explained as follows:

- As n falls, k rises marginally with an accompanying fall in the interest rate.

- However, the tax rate falls infinitesimally. The impact on welfare of both agents is infinitesimal as well.

Discount Factor In so far we have assumed that $\beta(= 0.96) > \theta(= 0.9)$. We now allow for these values to differ.

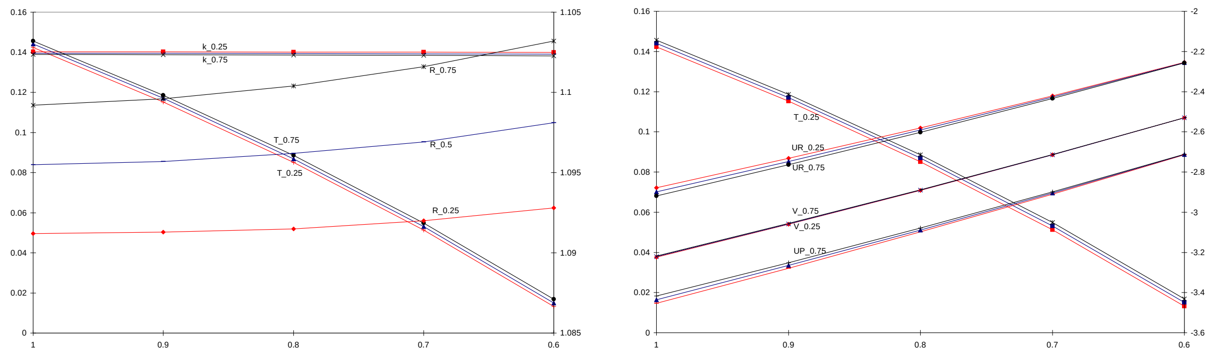


Figure 17: Impact of Change in Discount Factors (Same discount factor)

Firstly, we consider the case where the agents have the same discount factor δ. We initially assume that all the agents have perfect foresight such that $\delta = 1$ and then allow for myopia to set in. The agents then discount the future at a higher rate. When both agents discount the future at a higher rate, this leads to a decline in capital available per worker, initially at a slow rate but this rate eventually accelerates. This leads to a similar rise in the rate of interest. As before we note that the capital stock remains higher when a lower weight is attached to ω, that is, the less redistribution there is in the system.

We also note that, as the agents discount the future at a higher rate, this leads to a fall in the tax rate. The tax rate is initially around 14%, with the tax rate still higher under redistribution. However, as the agents discount the future at a higher rate, this leads to a fall in the tax rate such that when $\delta = 0.6$, the tax rate falls to between 0 and 2%. The lower tax rate compensates for some of the fall arising in steady state capital as a result of the higher discounting of the future.

The welfare levels of all the agents, as well as aggregate welfare, increase as the agents discount the future at a higher rate. We can explain the results as follows:

- As the agents discount the future at a higher rate, this has an adverse impact on the capital stock
- However, part of this is reversed as the tax rate falls as well
- The utility gains to the consumers is higher as a result of the higher level of consumption enjoyed as well as the lower tax rate combine to exceed the impact of the reduction in capital.

It can also be seen that as δ falls under 0.6, then the tax rate falls further and beyond a certain point, it is optimal to have no pensions policy in place.

In the second set of experiments, we allow for the rich to be "life-cyclers" such that they do not discount the future whilst allowing for the poor to

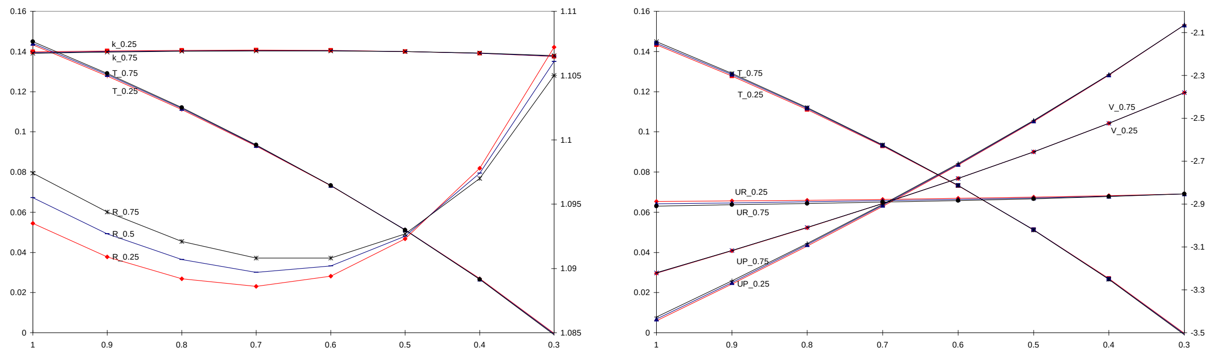


Figure 18: Impact of Change in Discount Factors (Differing Discount factors)

suffer from increasing levels of myopia such that $\beta \geqslant \theta$. As the poor discount the future at a higher rate, this has the effect of reducing the tax rate and this effect dominates the resulting reduction in capital as a result of lower saving, such that the capital stock initially rises.

In the early stages, we note that the tax rate is still higher the higher the weight attached to Pillar 1 but these results are overturned beyond a certain point. The tax rate keeps on falling as increasing levels of myopia set in. However, beyond a certain point, the resulting fall in capital from the decline in capital from the higher discounting of the poor results in a lower capital (despite the tax rate falling). At high levels of myopia, it is optimal to have no pensions at all.

The welfare of the rich improves as the poor discount the future at higher rates. This is in part due to the falling tax rate. The poor benefit from the fact that their consumption is higher. Beyond a certain point, the welfare of the poor is higher than that of the rich. Aggregate welfare rises as well. The results can be summarised as:

- Capital falls as poor discount at higher rate, but the tax rate falls as well;
- This initially leads to an increase in capital. However, beyond a certain point, even the fall in the tax rate cannot compensate for the fall in capital;
- The welfare of the rich improves as a result of lower taxes;
- The poor benefit from higher consumption.

5.5 An Application

We now test the model to a sample of 23 countries. The sample consists of 10 European countries (inc. Turkey); 6 countries in America (5 Latin America); 3 African countries and the others made up of Israel, Singapore, Australia and New Zealand. Though the World Bank introduced the 3 pillars

Country	n	Alpha	Beta	Pi	Psi
Arg	1.3	0.26	0.91	0.76	0.50
Auz	1.3	0.30	0.93	0.75	0.80
Blg	0.2	0.26	0.93	0.73	0.63
Brz	1.7	0.25	0.65	0.58	0.38
Chl	1.5	0.25	0.90	0.63	0.47
Egy	2.0	0.25	0.95	0.51	0.47
Fin	0.4	0.32	0.96	0.68	0.85
Fra	0.5	0.23	0.95	0.71	0.65
Ger	0.2	0.27	0.92	0.66	0.53
Ire	0.8	0.22	0.97	0.65	0.75
Isr	2.3	0.29	0.86	0.75	0.50
Ita	0.1	0.27	0.95	0.64	0.95
Mex	1.6	0.21	0.95	0.58	0.49
Neth	0.6	0.28	0.95	0.73	0.75
NZ	1.1	0.29	0.93	0.69	0.74
Nig	2.7	0.25	0.96	0.41	0.35
Per	1.8	0.29	0.88	0.76	0.62
Sng	2.3	0.22	0.95	0.74	0.45
SA	2.0	0.27	0.96	0.64	0.50
Swi	0.6	0.25	0.97	0.72	0.71
Tur	1.8	0.24	0.96	0.43	0.42
UK	0.3	0.22	0.96	0.75	0.72
US	1.0	0.27	0.95	0.77	0.65

Table 10: Parameter Summary

of pensions with the aim of reducing the stress on the first pillar due to an ageing population, one of the concerns raised (see Gill *et al.* (2004) and Boeri *et al.* (2006)) is that the first pillar still has a very high weight in the scheme and the third pillar is almost non-existent. In this section, we assume that $\omega = 0.75$, that is 75% of the pensions scheme is through the first pillar. We estimate probable values for α, β, π, ψ, and n as before and assume that the agents have the same discount factor. Agents thus differ according to their productivity only. An overview of the parameters is given in Table 10.

Country	k	R	Tau	UR	UP	V
Arg	0.141	1.105	0.116	-2.814	-4.006	-3.100
Auz	0.169	1.042	0.044	-3.022	-3.438	-3.126
Blg	0.147	1.077	0.119	-2.843	-3.650	-3.061
Brz	0.114	1.271	0.038	-2.342	-3.893	-2.993
Chl	0.128	1.170	0.123	-2.759	-4.056	-3.239
Egy	0.120	1.227	0.123	-2.833	-4.179	-3.492
Fin	0.179	1.031	0.009	-3.165	-3.481	-3.266
Fra	0.136	1.067	0.179	-2.739	-3.460	-2.948
Ger	0.142	1.121	0.091	-2.877	-4.009	-3.262
Ire	0.134	1.056	0.202	-2.717	-3.197	-2.885
Isr	0.148	1.127	0.041	-2.863	-4.107	-3.174
Ita	0.164	1.010	0.114	-2.912	-3.003	-2.945
Mex	0.114	1.167	0.206	-2.664	-3.840	-3.158
Neth	0.160	1.046	0.088	-2.963	-3.486	-3.104
NZ	0.161	1.061	0.060	-2.977	-3.532	-3.149
Nig	0.101	1.401	0.108	-2.828	-4.750	-3.962
Per	0.156	1.082	0.047	-2.896	-3.760	-3.103
Sng	0.123	1.130	0.199	-2.695	-3.985	-3.031
SA	0.136	1.161	0.100	-2.937	-4.190	-3.388
Swi	0.147	1.054	0.150	-2.857	-3.453	-3.024
Tur	0.109	1.297	0.134	-2.802	-4.356	-3.688
UK	0.136	1.042	0.202	-2.701	-3.242	-2.836
US	0.153	1.063	0.108	-2.918	-3.684	-3.094

Table 11: Results Summary for omega=0.75

Based on the above parameters, we can derive the optimal tax rate for the countries and also consider the implications for the other endogenous parameters. The results for $\omega = 0.75$ are shown in Table 11.

We find that for $\omega = 0.75$, the tax rate varies between 0.9% for Finland and 20.6% for Mexico with an average of 11.3% for the sample. All the economies are dynamically efficient for the given parameters with the interest rate being 1% for Italy and 40% for Nigeria with an average of 12% for the sample. Allowing ω to fall, that is putting a higher weight on Pillar 2 increases capital and reduces the interest rate; the tax rate falls marginally and the welfare of the rich increase but the welfare of the poor and aggregate welfare falls.

5.6 Conclusions

The issue of pensions and pensions reforms is likely to remain on the policy agenda for a while. Given the economic and political ramifications it entails, the design of an optimal pensions policy has to take into account the various parameters in the economy and the extent of inequality. In this paper, we have formalised the intuition behind the WB model and designed a pensions system based on 3 pillars with the triple objectives of achieving redistribution, consumption-smoothing and promoting voluntary savings. We extend a pay-as-you-go pensions system to include the two pillars as suggested by the World Bank (1994). One of our key results suggests that the weight the planner attaches to Pillar 1 or Pillar 2 will be influenced by his inequality aversion. However, this does not have a significant impact on the tax rate. A more redistributive pensions system hinders capital accumulation. The poor will prefer a higher weight on Pillar 1 due to its redistributive nature. On the other hand, the rich will prefer Pillar 2 since it is closely related to ability and contribution rates. Our other results suggest that a richer economy is able to afford a higher level of pensions in that as the proportion of rich increases, the tax rate increases. Similarly, the smaller the difference in the productivity of the rich and the poor, the higher the tax rate. One of the key policy findings emanating from our model is that as the population ages, the tax rate is not affected to a large extent. This has important implications for fiscal policy in face of an ageing population. However, this result needs to be qualified in that there is no debt in our model and the pensions payment is allowed to vary.

Though this model had been kept deliberately simple for analytical tractability, there are two natural extensions that could follow. Including a fully-funded second pillar would bring the model closer to the World Bank's (1994) initial proposal, however, as we have mentioned the private sector should not necessarily yield a higher return than the public sector once the administrative costs and risks are considered (Barr, 2001). The other extension would be to include the three pillars as proposed by the World Bank (2005) and include Pillar 0 and Pillar 4. This would include an element of means-tested pensions (Pillar 0) and include an element of altruism in the model to account for the informal forms of pensions (Pillar 4).

Appendix 5.1

The optimal level of consumption and savings of the rich is given as:

$$
\begin{aligned}
c^{y,r} &= \frac{1}{[1+\beta]}\left[w_t(1-\tau)+\frac{[\omega\,(1-\pi)\,(\psi-1)+1]\,(1+n)\,\tau w_{t+1}}{R_{t+1}}\right] \\
c^{o,r} &= \frac{\beta R_{t+1}}{[1+\beta]}\left[w_t(1-\tau)+\frac{[\omega\,(1-\pi)\,(\psi-1)+1]\,(1+n)\,\tau w_{t+1}}{R_{t+1}}\right] \\
s^{y,r} &= \frac{1}{[1+\beta]}\left[\beta w_t(1-\tau)-\frac{[\omega\,(1-\pi)\,(\psi-1)+1]\,(1+n)\,\tau w_{t+1}}{R_{t+1}}\right]
\end{aligned}
$$

That of the poor is as follows:

$$
\begin{aligned}
c^{y,p} &= \frac{1}{[1+\theta]}\left[\psi w_t(1-\tau)+\frac{[\omega\pi\,(1-\psi)+\psi]\,(1+n)\,\tau w_{t+1}}{R_{t+1}}\right] \\
c^{o,p} &= \frac{\theta R_{t+1}}{[1+\theta]}\left[\psi w_t(1-\tau)+\frac{[\omega\pi\,(1-\psi)+\psi]\,(1+n)\,\tau w_{t+1}}{R_{t+1}}\right] \\
s^{y,p} &= \frac{1}{[1+\theta]}\left[\theta\psi w_t(1-\tau)-\frac{[\omega\pi\,(1-\psi)+\psi]\,(1+n)\,\tau w_{t+1}}{R_{t+1}}\right]
\end{aligned}
$$

The steady-state capital is given by:

$$
\left(k^* = \left[\frac{\{\beta(1+\theta)\pi+(1+\beta)\theta(1-\pi)\psi\}\alpha(1-\alpha)(1-\tau)}{(1+n)\{(1+\beta)(1+\theta)\alpha+\{(1+\theta)\pi[\omega(1-\pi)(\psi-1)+1]+(1+\beta)(1-\pi)[\omega\pi(1-\psi)+\psi]\}(1-\alpha)\tau\}}\right]^{\frac{1}{1-\alpha}}\right) \tag{1}
$$

In intensive form, the welfare function can be expressed as

$$V = \pi\left[\ln(c^{y,r}) + \beta\ln(c^{o,r})\right] + (1-\pi)\left[\ln(c^{y,p}) + \theta\ln(c^{o,p})\right] \qquad (2)$$

We can then express V as a function of the policy-parameters ω and τ and non-policy parameters $(\alpha, \beta, \theta, \pi, \psi, n)$

$$V = U\ln\left[J + [L + M\omega]\,\tau\right] + Z\ln\left[N + [O + P\omega]\,\tau\right] + Q\left[\ln A(1-\tau) - \ln\left[B + \{G + H\omega\}\,\tau\right]\right]$$

where A, B, G, H, J, Q, U, L, M and Z are all non-policy parameters.

and defined as:

$A = (\beta * (1+\theta) * \pi + (1+\beta) * \theta * (1-\pi) * \psi) * \alpha * (1-\alpha)$

$B = (1+\beta) * (1+\theta) * (1+n) * \alpha$

$C = (1+\theta) * \pi * (1+n) * (1-\alpha)$

$D = (1-\pi) * (\psi - 1)$

$E = (1+\beta) * (1-\pi) * (1+n) * (1-\alpha)$

$F = \pi * (1-\psi)$

$G = C + E * \psi$

$H = (E * F + C * D)$

$J = \alpha * B$

$L = \alpha * G + A * (1+n)$

$$M = \alpha * H + A * (1+n) * (1-\pi) * (\psi - 1)$$

$$N = B * \alpha * \psi$$

$$O = G * \alpha * \psi + A * (1+n) * \psi$$

$$P = H * \alpha * \psi + A * (1+n) * \pi * (1-\psi)$$

$$Q = \left[\frac{(\pi*((1+\beta)*\alpha-(\alpha-1))+(1-\pi)*((1+\theta)*\alpha-(\alpha-1)))}{(1-\alpha)}\right]$$

$$U = (1+\beta) * \pi$$

$$Z = (1-\pi) * (1+\theta)$$

Proof of Proposition 12 (Based on Eqn. (21)**)**

$$(1a) : \frac{\partial k}{\partial \delta} = \frac{1}{1-\alpha}\left[\frac{\{\alpha+[\pi+(1-\pi)\psi](1-\alpha)\tau\}[\{\pi+(1-\pi)\psi\}\alpha(1-\alpha)(1-\tau)]}{(1+n)[\{(1+\delta)\alpha+[\pi+(1-\pi)\psi](1-\alpha)\tau\}]^2}\right]^{\frac{\alpha}{1-\alpha}} > 0$$

$$(1b) : \frac{\partial k}{\partial \pi} = \frac{1}{1-\alpha}\left[\frac{\delta(1+\delta)(1-\alpha)\alpha^2(1-\psi)(1-\tau)}{(1+n)[\{(1+\delta)\alpha+[\pi+(1-\pi)\psi](1-\alpha)\tau\}]^2}\right]^{\frac{\alpha}{1-\alpha}} > 0$$

$$(1c) : \frac{\partial k}{\partial \psi} = \frac{1}{1-\alpha}\left[\frac{\delta(1+\delta)(1-\alpha)\alpha^2(1-\pi)(1-\tau)}{(1+n)[\{(1+\delta)\alpha+[\pi+(1-\pi)\psi](1-\alpha)\tau\}]^2}\right]^{\frac{\alpha}{1-\alpha}} > 0$$

$$(1d) : \frac{\partial k}{\partial \tau} = \frac{1}{1-\alpha}\left[\frac{-[\delta\alpha(1-\alpha)\{\pi+(1-\pi)\psi\}[(1+\delta)\alpha+(1-\alpha)\{\pi+(1-\pi)\psi\}]]}{(1+n)[\{(1+\delta)\alpha+[\pi+(1-\pi)\psi](1-\alpha)\tau\}]^2}\right]^{\frac{\alpha}{1-\alpha}} < 0$$

$$(1e) : \frac{\partial k}{\partial n} = \frac{1}{1-\alpha}\left[\frac{-[\delta\{\pi+(1-\pi)\psi\}\alpha(1-\alpha)(1-\tau)\{(1+\delta)\alpha+[\pi+(1-\pi)\psi](1-\alpha)\tau\}]}{[(1+n)\{(1+\delta)\alpha+[\pi+(1-\pi)\psi](1-\alpha)\tau\}]^2}\right]^{\frac{\alpha}{1-\alpha}} < 0$$

Proof of Proposition 13 (Based on Eqn. (1)**)**

$$\begin{aligned}
\frac{\partial k}{\partial \omega} &= \frac{1}{1-\alpha}\left[\frac{-AE}{\{B+\omega E\}^2}\right]^{\frac{\alpha}{1-\alpha}} < 0 \text{ for } \beta > \theta \\
\frac{\partial k}{\partial \omega} &= \frac{1}{1-\alpha}\left[\frac{-AE}{\{B+\omega E\}^2}\right]^{\frac{\alpha}{1-\alpha}} = 0 \text{ for } \beta = \theta
\end{aligned}$$

where :

$$A = \{\beta\,(1+\theta)\,\pi + (1+\beta)\,\theta\,(1-\pi)\,\psi\}\,\alpha(1-\alpha)(1-\tau) > 0$$

$$B = (1+n)\,[(1+\beta)\,(1+\theta)\,\alpha + \{[(1+\theta)\,\pi + \psi\,(1+\beta)\,(1-\pi)]\,(1-\alpha)\tau\}] > 0$$

$$E = [\beta-\theta]\,(1-\psi)\,\pi\,(1-\pi)\,(1+n)\,(1-\alpha)\tau \substack{>0 \text{ for } \beta>\theta \\ =0 \text{ for } \beta=\theta}$$

Since we are only concerned with the case of $\beta > \theta$ and $\beta = \theta$, an increase in ω can only lead to a reduction in k or have no impact on k, depending on β and θ.

Proof of Proposition 14

$$(3a)\quad : \quad \frac{\partial V}{\partial \pi} = \left[\frac{(1+\delta)\,\alpha - \delta\,(1-\alpha)}{(1-\alpha)}\right]\left[\frac{(1-\psi)}{[\pi + (1-\pi)\,\psi]}\right] - (1+\delta)\ln\psi > 0$$

$$(3b)\quad : \quad \frac{\partial UR}{\partial \pi} = \frac{\partial UP}{\partial \pi} = \left[\frac{(1+\delta)\,\alpha + \delta\,(\alpha-1)}{(1-\alpha)}\right]\left[\frac{(1-\psi)}{[\pi + (1-\pi)\,\psi]}\right] > 0$$

Proof of Proposition 15

$$(4a)\quad : \quad \frac{\partial V}{\partial \psi} = \left[\frac{(1+\delta)\,\alpha - \delta\,(1-\alpha)}{(1-\alpha)}\right]\left[\frac{(1-\pi)}{[\pi + (1-\pi)\,\psi]}\right] + \left[\frac{(1+\delta)\,(1-\pi)}{\psi}\right] > 0$$

$$(4b)\quad : \quad \frac{\partial UR}{\partial \psi} = \left[\frac{(1+\delta)\,\alpha + \delta\,(\alpha-1)}{(1-\alpha)}\right]\left[\frac{(1-\pi)}{[\pi + (1-\pi)\,\psi]}\right] > 0$$

$$(4c)\quad : \quad \frac{\partial UP}{\partial \psi} = \left[\frac{(1+\delta)\,\alpha + \delta\,(\alpha-1)}{(1-\alpha)}\right]\left[\frac{(1-\pi)}{[\pi + (1-\pi)\,\psi]}\right] + \left[\frac{(1+\delta)}{\psi}\right] > 0$$

Proof of Proposition 16

$$(5) : \frac{\partial V}{\partial n} = \frac{\partial UR}{\partial n} = \frac{\partial UP}{\partial n} = -\left[\frac{(1+\delta)\,\alpha - \delta\,(1-\alpha)}{(1-\alpha)\,(1+n)}\right] < 0$$

Proof to Proposition 17

$$(6) \quad : \quad \frac{\partial UR}{\partial \omega} = \left[\frac{-a\,(1-\psi)\,(1+\delta)\,(1-\pi)\,(1+n)\,\tau}{[\alpha\,[b+c\tau] + a\,[\omega\,(1-\pi)\,(\psi-1)+1]\,(1+n)\,\tau]} \right] < 0$$

$$(6) \quad : \quad \frac{\partial UP}{\partial \omega} = \left[\frac{a\,(1+\delta)\,\pi\,(1-\psi)\,(1+n)\,\tau}{[\alpha\psi\,[b+c\tau] + a\,[\omega\pi\,(1-\psi)+\psi]\,(1+n)\,\tau]} \right] > 0$$

Proof of Proposition 18

$$(7) \quad : \quad \frac{\partial UR}{\partial \tau} = \left[\begin{array}{c} -\left[\frac{(1+\delta)\alpha-(\alpha-1)}{(1-\alpha)}\right]\left[\frac{1}{(1-\tau)}+\frac{c}{(b+c\tau)}\right] \\ +\left[\frac{(1+\delta)[\alpha c+a[\omega(1-\pi)(\psi-1)+1](1+n)]}{[\alpha[b+c\tau]+a[\omega(1-\pi)(\psi-1)+1](1+n)\tau]}\right] \end{array} \right] = -A2 + B2$$

$$(7) \quad : \quad \frac{\partial UP}{\partial \tau} = \left[\begin{array}{c} -\left[\frac{(1+\delta)\alpha-(\alpha-1)}{(1-\alpha)}\right]\left[\frac{1}{(1-\tau)}+\frac{c}{(b+c\tau)}\right] \\ +\left[\frac{(1+\delta)[\alpha\psi c+a[\omega\pi(1-\psi)+\psi](1+n)]}{[\alpha\psi[b+c\tau]+a[\omega\pi(1-\psi)+\psi](1+n)\tau]}\right] \end{array} \right] = -A2 + B3$$

where:

$a = \delta\,\{\pi + (1-\pi)\,\psi\}\,\alpha(1-\alpha);$

$b = (1+\delta)\,(1+n)\,\alpha;$

$c = [\pi + (1-\pi)\,\psi]\,(1+n)\,(1-\alpha).$

From the above, the utility of the rich increases if $B2 > -A2$. On the other hand, the utility of the poor increases if $B3 > -A2$. UR and UP will differ if $B2 \neq B3$. If $B3 > B2$, then $\frac{\partial UP}{\partial \tau} > \frac{\partial UR}{\partial \tau}$.

6 CONCLUSIONS AND FUTURE EXTENSIONS

6.1 Conclusions

Whilts the issue of pensions reforms has temporarily taken the backseat due to the global crisis, it is bound to be back in the limelight due to the financial and economic implications an ageing population entails. One of the crucial issues policymakers will have to deal with is ensure how the elderly can be catered for without jeopardizing financial and fiscal sustainability. As things stand, without reforms, many of the PAYG pensions systems are expected to go bust. The problems have been compounded by the economic crisis which has caused the value of many pensions funds to fall, or in some cases collapse. The fiscal stimulus put in place in the wake of the crisis has also weakened the fiscal positions of many economies.

This thesis has considered the ageing problem and considered how best the pensions systems can be reformed to maximise aggregate social welfare. We developed a standard two-period OLG model in a general equilibrium framework with heterogeneous agents, profit-maximising firms and a benevolent social welfare maximiser, and considered how the pensions system in place affects the level of welfare. We considered which of the PAYG and FF schemes are optimal, or whether a hybrid system with a combination of both the PAYG and FF might be most appropriate. In particular, we considered the consumption-smoothing and redistributive role of pensions by investigating whether pensions should be ability-based or should there be a transfer from the rich to the poor. We considered a combination of redistributive mechanisms: in some cases, only the rich fund for the pensions whilst it is optional for the poor; whilst the poor always benefit when a pensions system is in place, this is not necessarily the case for the rich.

Our key result, consistent with Samuelson (1975) and Feldstein (1985), suggests that the pensions system in place has major implications for capital formation. The pensions system determines capital accumulation, which in turn has implications for output, the interest rate, wages, consumption and by extension welfare. Compared to a PAYG system, capital is always higher under a FF system due to the crowding out induced under the PAYG. Feldstein (1985), had shown the existence of an optimal PAYG pensions system in a dynamically efficient economy. By extending his model to a general equilibrium framework, we were able to show that a PAYG yields a much lower welfare than the FF scheme. Consistent with Aaron (1966), we also find that in so far as an economy is dynamically efficient, a FF system is optimal for consumption smoothing purposes. We find that there is a higher possibility of dynamic inefficiency arising under a FF as there is a higher potential for over accumulation of capital. Dynamic efficiency can be restored in a dynamically inefficient economy by introducing a PAYG pensions system, which leads to a reduction in capital. However, one of our key findings departs from Aaron, in that there is room for a PAYG pensions system in a dynamically efficient economy that is approaching dynamic inefficiency, that is, there is a potential for over-accumulation of capital. The introduction of a PAYG in such an economy will ensure the economy remains dynamically efficient. This results needs to be qualified in that it only holds for a small range of parameter values when the economy is approaching dynamic inefficiency. We also find that in the models where the agents differ in terms of productivity, richer economies—characterised by a higher proportion of high productivity agents—tend to have a higher level of capital and welfare.

In the heterogeneous framework, we considered a combination of the redistributive or consumption-smoothing roles of pensions. The pensions could be FF or PAYG, or there could be a hybrid system in place. We find that in a dynamically efficient economy, by and large, a FF scheme is optimal, except for the case when there is a potential for capital over-accumulation. When rich economies over-accumulate capital, a PAYG pensions system coupled with intragenerational redistribution can lead the economy to dynamic efficiency. However, when redistribution is effected through pensions, it has an adverse impact on capital formation when the poor discount the future at a higher rate. The poor always prefer a higher level of redistribution than the rich, although in our framework, redistribution always leads to higher aggregate welfare. Our other results suggest that a rich economy require less redistribution but can afford to be more generous. We also find that there are ranges of parameter values for which the role of pensions is best restricted to consumption smoothing and redistribution is best effected through instruments other than pensions.

In line with Conde-Ruiz and Galasso (2005), we also considered both inter and intragenerational redistribution simultaneously, with an element of inequality aversion from the planner. We find that when the agents have the same discount factor, the tax rate and the timing of redistribution tilt in such a way to ensure that capital does not change significantly and the economy remains dynamically efficient at all times. With higher inequality, the tax rate is higher and intragenerational transfers are preferred. Richer economies require less redistribution but can afford intergenerational transfers. The tilting in the timing of redistribution mitigates the adverse impact of the higher tax rate on capital formation. We also find that as population ages, intergenerational redistribution is preferred since it counters the increase in capital per worker and ensures the economy remains dynamically efficient. Both inter and intra generational redistribution are supported only within a range; outside this range only one instrument is optimal, that is, intragenerational redistribution for poor economies and intergenerational redistribution for rich economies.

In line with the World Bank's "Averting the Old Age Crisis" (1994), we formalised the intuition behind the World Bank model and designed a pensions system with three pillars. Pillar 1 is redistributive, Pillar 2 is ability based and Pillar 3 is optional. The results are consistent with our previous findings whereby the higher the redistributive component of the pensions system, the higher (lower) the welfare of the poor (rich) but aggregate welfare increases. However, whether the planner chooses Pillar 1 or Pillar 2 makes a marginal difference to the optimal tax rate.

While we have discussed the policy issues in the individual chapters, some of the main ones need highlighting. As discussed, the issue of pensions reforms is likely to remain on the policy agenda for the foreseeable future. In this regard, it is important that the reforms, whether parametric or wholesale, are implemented in a way to ensure economic and social efficiency whilst considering long-term sustainability. One of our key findings suggests that in a dynamically efficient economy, the FF scheme is optimal. However, the design of the pensions system should also take into account the whole range of redistributive issues associated with such a reform, especially when transiting from a PAYG to a FF system. The need to redistribute to tackle inequality and poverty as well as ensuring as broad a coverage of the social safety nets also need attention. These are some of the key issues that future research needs to investigate further.

6.2 Future Extensions

To ensure analytical tractability and ease of manipulation, we have had recourse to a range of simplifying assumptions. The key assumptions driving our results pertain to the loglinear utility function and Cobb-Douglas production function. Agents live for two-periods, with certainty, and there are no bequests. In line with Becker (1990), the poor discount the future at a higher rate than the rich. The other assumptions imply that there is no adverse impact of pensions on labour supply since one unit of labour is provided inelastically. Moreover, we did not consider the transition cost of moving from a PAYG to a FF system, which might be significant. Implicitly, we assumed full coverage of pensions and all the economic activities took place in the formal sector. We also assumed that the switch in the pensions system would be based solely on economic considerations and there were no political economy issues involved. Whilst some of the results, especially those pertaining to capital formation, will continue to hold under different frameworks, those on welfare will change depending on the framework used. We next consider how some of the assumptions can be relaxed and how the research can be extended further.

Utility Functions

The loglinear utility function, coupled with the Cobb Douglas production function, ensure the existence of the competitive equilibrium as defined. Given the nature of the OLG models, analytical tractability is lost very quickly with other general functions. The loglinear utility function is one way of keeping the model manageable (Balasko and Shell, 1981). However, these assumptions could be relaxed by introducing more general utility functions. The main way in which it would affect our existing result is that in the current set up the transfer from the rich to the poor always increases aggregate welfare because the utility of the poor increases by more than the fall in that of the rich due to the concave function. However, this would not be the case with convex functions.

Ageing through increased Longevity

In this thesis, we have considered ageing through an exogenous population growth rate, whereby a fall in the population growth rate leads to an increase in the dependency ratio, or ratio of retirees to workers. However, another dimension through which population ageing is taking place is through increased longevity. In our framework, agents live for two periods with certainty. One could consider adding the longevity element a la Blanchard (1985) continuous-time model with age invariant mortality rates.

More general demographic structures have also been considered by Lau (2009). Bringing in a risk of mortality outside the specific periods in the Diamond spirit will extend the model and ageing process in another dimension. Whilst we have considered one period as one year, this could be easily extended to longer periods with one generation being 20-30 years, as in Feldstein (1984). Other models have gone beyond the traditional two-period (see de la Croix and Michel, 2002), and included additional finite periods (Auerbach and Kotlikoff, 1987), with a probability of survival into the additional periods.

Impact of Social Security on Labour Supply

We have also assumed that agents supply one unit of labour inelastically and their choice is not affected by the provision of pensions. However, there is evidence (Hu, 1979) that if transfers are tied to the retirement decision, the pensions system introduces distortions into the labour supply choice. This is something we have not considered in our model. Introducing this element will increase the cost of having a PAYG pensions system beyond the crowding out of capital we have considered.

Fertility and Bequests

A related element pertains to the exogenous population growth rate. Barro and Becker (1989) and Becker and Mulligan (1997) have considered how fertility choice could be endogenised and introducing this in an OLG model with pensions could provide for a richer analytical environment. Staying within the structure of the model, the agents are not altruists and do not leave any bequests for the next generation. However, bequests, whether altruistic or accidental, are important (Becker, 1974; Abel, 1985). Introducing bequests would provide another dimension in which the environment could be enriched. What bequests will do is mitigate the crowding out induced by the PAYG pensions system. Hence, we could have additional interactions with implications for capital formation whereby the reduction in labour supply and the bequests would increase capital whereas the increased fertility will dilute capital. The overall impact will of course depend on which of the effects dominate.

Steady State, Transition Costs, Political Economy Issues

The current analysis has been carried out in steady state. What we have not considered is how the transition from a PAYG to a FF system takes place and how to deal with the associated transition cost, which is an important and highly debated issue. Funding the transition cost, either by accumulating debt or using the proceeds of privatization, has attracted significant attention since the pensions reforms in Latin America and "Averting the Old Age Crisis". Notwithstanding the economic ramifications, there are significant political economy issues at play (see for example, Conde Ruiz and Galasso, 2005, Galasso and Profetta, 2002, 2004). These political economy issues are one of the major factors hindering the reforms of pensions systems, even unsustainable ones. Further research in the political economy issues will shed light on why pensions systems are delayed.

Market Imperfections

Finally, the economic environment could be enriched further. Whilst we have assumed perfect capital markets, the real world is characterised by market imperfections. For instance, households could be credit constrained and this will hinder their ability to borrow. Markets may not always clear such that there can be an excess of labour supply thereby leading to unemployment. Moreover, we have assumed there is no informal sector and everyone is covered by the pensions system in place. However, evidence suggests that the issue of coverage is important, especially in developing countries (Gill et al., 2004). More recently, the financial crisis has shown that there are previously unforeseen risks that the FF system might be subject to and this can affect the value of pensions households will receive.

We have used a simple framework to demonstrate which pensions system is optimal and maximises welfare per head. One of our key results has shown that a hybrid system with both a FF and a PAYG element, with the latter catering for redistribution, is optimal. However, there are various other issues which we have not considered, but highlighted above. Taking into account those factors, while undeniably increasing the complexity and tractability of the model, could provide additional insights into pensions reforms.

Bibliography

1. Aaron, H. (1966), "The Social Security Paradox", *Canadian Journal of Economics and Political Science,* 32, 371-376.

2. Abel, A., (1985), "Precautionary Saving and Accidental Bequest," *American Economic Review*, 75(4):777-791"

3. Alesina A., Rodrik, D.,(1994), "Distributive Politics and Economic Growth", *The Quarterly Journal of Economics*, 109(2), 465-490.

4. Alesina A., Perotti, R., (1996), "Income Distribution, Political Instability, and Investment", *European Economic Review*, 40(6), 1203-1228.

5. Antolin, P., de Serres, A., and de la Maisonneuve, C., (2004). "Long-Term Budgetary Implications of Tax-favoured Retirement Saving Plans." *OECD Working Paper 393*, Economic Department, OECD, Paris.

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

7. Bailliu, J., and Reisen H., (1997). "Do Funded Pensions Contribute to Higher Aggregate Savings? A Cross-Country Analysis." In *Pensions, Savings, and Capital Flows: From Ageing to Emerging Markets*, ed. Helmut Reisen, 113–31. Paris: Organisation for Economic Co-operation and Development.

8. Balasko Y., Shell., K., (1981), "The Overlapping-Generations Model. III. The Case of Log-Linear Utility Functions," *Journal of Economic Theory*, 24:143-152

9. Barr, N., (2001). "The Truth about Pension Reform." *Finance and Development* 38 (3): 6–9.

10. Barro, R. J., & Becker, G.S., (1989), "Fertility Choice in a Model of Economic Growth," *Econometrica*, 57(2): 481-501

11. Becker, G.S., (1974), "A Theory of Social Interactions", *Journal of Political Economy*, 82(6):1063-1091

12. Becker, G.S., (1990), "Optimal Discounting of the Future," Department of Economics, University of Chicago.

13. Becker, G. S., & Mulligan, C. B., (1997), "The Endogenous Deter-

mination of Time Preference," *The Quarterly Journal of Economics*, 112(3):729-58

14. Bertranou, F., van Ginneken, W., and Solorio, C., (2004) "The Impact of Tax-Financed Pensions on Poverty Reduction in Latin America: Evidence from Argentina, Brazil, Chile, Costa Rica, and Uruguay." *International Social Security Review* 57 (4): 3–18.

15. Blanchard, O., (1985), "Debt, Deficits and Finite Horizons," *Journal of Political Economy*, 93:223-247

16. Boeri, T., Bovenberg, L., Coeuré, B. and Roberts, A. (2006)*, 'Dealing with the new giants: Rethinking the role of pension funds.',* Geneva Reports on the World Economy *8.*

17. Bohn, H., (1999), "Should the U.S. Social Security Trust Hold Equities? An Intergenerational Welfare Analysis", *Review of Economic Dynamics*, 2(3), 666-97.

18. Bossi, L., (2007), "Intergenerational Risk Shifting through Social Security and Bailout Politics", *Journal of Economic Dynamics and Control* , doi:10.1016/j.jedc.2007.10.001

19. Bosworth, B., and Burtless G., (2003). "*Pension Reform and Saving*" Brookings Institution, Washington, DC.

20. Conesa, J.C., Krueger, D., (1999), "Social Security Reform with Heterogeneous Agents", *Review of Economic Dynamics,* 2(4), 757-95.

21. Conde-Ruiz, J.I., Galasso, V., (2005), "Positive Arithmetic of the Welfare State", *Journal of Public Economics,* 89, 933-55.

22. Corbo,V., and Schmidt-Hebbel, K.,. (2003) "Macroeconomic Effects of the Pension Reform in Chile." In *Pension Reforms: Results and Challenges*, ed. International Federation of Pension Fund Administrators, 241–329. Santiago, Chile: International Federation of Pension Fund Administrators.

23. De la Croix, D., and Michel, P., (2002): *A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations*, Cambridge University Press, Cambridge

24. Diamond, P., (1965), "National Debt in a Neoclassical Growth Model", *American Economic Review*, 55(5), 1126-1150.

25. Edwards, S., and Alejandra Cox-Edwards. (2002) "Social Security Pri-

vatization Reform and Labor Markets: The Case of Chile." *NBER Working Paper 8924*, National Bureau of Economic Research, Cambridge, MA.

26. Feldstein, M. (1974), "Social Security, Induced Retirement, and Aggregate Capital Accumulation", *Journal of Political Economy*, 82(5), 905-26.

27. Feldstein, M. (1985), "The Optimal Level of Social Security Benefits", *Quarterly Journal of Economics*, 10(2), 303-320.

28. Feldstein, M., (1996), "Social Security and Saving: New Time Series Evidence", *National Tax Journal*, Vol. 49 (2), pp. 151–64.

29. Feldstein, M., (2008) "Rethinking Social Insurance", American Economic Association Presidential Address

30. Friedman, M., (1962), *Capitalism and Freedom*, Chicago: The University of Chicago Press

31. Galasso, V., Profeta, P., (2002), "Political Economy Models of Social Security: A Survey", *European Journal of Political Economy*, 18, 1–29.

32. Galasso, V., Profeta, P., (2004), "Lessons for an Ageing Society: The

Political Sustainability of Social Security Systems", *Economic Policy,* 19, 63–115.

33. Galasso, V., Profeta, P., (2007),. "How does Ageing Affect the Welfare State?", *European Journal of Political Economy*, 23, 554–563

34. Galor, O., (1992), "The Choice of Factor Mobility in a Dynamic World", *Journal of Population Economics*, 5(2), 135-44.

35. Galor, O., Polemarchakis, H., (1987), "Intertemporal Equilibrium and the Transfer Paradox", *Review of Economic Studies*, 54(1), 147-56.

36. Gill, I., Packard, T., and Yermo, J., (2004). *Keeping the Promise of Old-Age Income Security in Latin America.* Washington, DC: World Bank.

37. Gillion C., Turner, J., Bailey, C., Latulippe, D., (Eds.) (2000): *Social Security Pensions: Development and Reform*, International Labour Office, Geneva

38. Hansson I., Stuart C., (1989), "Social Security as Trade among Living Generations", *American Economic Review*, 79(5), 1182-1195.

39. Haveman, R., (1988): *Starting Even: An Equal Opportunity Program to Combat the Nation's New Poverty*, Simon and Schuster, New York

40. Holzmann, R., (1997). "Pension Reform, Financial Market Development, and Economic Growth: Preliminary Evidence from Chile." IMF Staff Papers 44 (2): 149–78.

41. International Labour Organisation, *LABORSTA - ILO Database on Labour Statistics,* ILO Bureau of Statistics: http://laborsta.ilo.org/

42. International Monetary Fund, *International Financial Statistics*, Online Database

43. James, E., Cox-Edwards, A., and Wong.R., (2003) "The Gender Impact of Pension Reform." *Journal of Pension Economics and Finance* 2 (2): 181–219.

44. Krueger, D., and Kubler, F., (2006), "Pareto Improving Social Security Reform when Financial Markets are Incomplete, *American Economic Review*, 96(3), 737-755.

45. Kuznets, S., (1955), "Economic Growth and Income Inequality", *American Economic Review*, 45(1), 1-28.

46. Lau, S.H.P., (2009), “Demographic Structure and Capital Accumulation: A Quantitative Assessment,” *Journal of Economic Dynamics and Control*, 33:554-567

47. Lerner, A.P., (1959) “Consumption-Loan Interest and Money”, *The Journal of Political Economy*, Vol. 67(5), 512-518.

48. Nehru, V., Dhareshwar, A., (1995), "A New Database on Physical Capital Stock: Sources, Methodology and Results", World Bank, Washington, DC.

49. Packard, T., (2002). “Is There a Positive Incentive Effect from Privatization Social Security? Panel Evidence from Latin America.” *Journal of Pensions Economics and Finance* 1 (2, July): 89–109.

50. Persson, T., Tabellini, G., (1994), “Is Inequality Harmful for Growth?”, *American Economic Review*, 84(3), 600-621.

51. Persson, T., Tabellini G., (2000): *Political Economics: Explaining Economic Policy*, MIT Press

52. Pigou A.C.,(1920): *The Economics of Welfare*, Macmillan, London

53. Poterba, J., (2001), “Demographic Structure and Asset Returns”, *Review of Economics and Statistics*, Vol. 83(4), pp. 565-584.

54. Razin A.,Sadka, E., Swagel, P., (2002), "The Aging Population and the Size of the Welfare State”, *Journal of Political Economy*, 110(4), 900-918

55. Samuelson, P.A. (1958), “An Exact Consumption-Loan Model of Interest with or without the Contrivance of Money”, *Journal of Political Economy*, 66(6), 467-482.

56. Samuelson, P.A. (1975), “Optimum Social Security in a Life-Cycle Growth Model”, *International Economic Review,* 16(3), 539-544.

57. Shiller, J., (1999), “Social Security and Institutions for Intergenerational, Intragenerational, and International Risk Sharing”, *Carnegie Rochester Conference on Public Policy,* 50, 165-204

58. Smith, A., (1776), *An Inquiry into the Nature and the Causes of the Wealth of Nations* (http://www.adamsmith.org/smith/won-index.htm)

59. United Nations (1999), *The World at Six Billion*, New York, USA

60. Vittas, D. (2000) “Pension Reform and Capital Market Development:

'Feasibility' and 'Impact' Preconditions." *Policy Research Working Paper 2414*, Development Research Group, Finance, World Bank, Washington, DC.

61. Walker, E., and Lefort, F., (2001) "Pension Reforms and Capital Markets: Are There Any (Hard) Links?" *Social Protection Discussion Paper 0201*, World Bank, Washington, DC.

62. World Bank. 1994. *Averting the Old-Age Crisis: Policies to Protect the Old and Promote Growth.* New York: Oxford University Press.

63. World Bank. 2005. *Old-Age Income Support in the Twenty-first Century: An International Perspective on Pension Systems and Reform,* Web version

64. World Bank, *World Development Indicators,* Online Database

65. Yermo, J., (2002) "The Performance of Funded Pension Systems in Latin America." Paper prepared for the Office of the Chief Economist, LAC Regional Study on Social Security Reform, World Bank, Washington, DC.