The international platform of Ludwig-Maximilians University's Center for Economic Studies and the Ifo Institute

CESifo Working Papers

www.cesifo.org/wp

The Intergenerational Welfare State

Torben M. Andersen
Joydeep Bhattacharya

CESIFO WORKING PAPER NO. 4359
CATEGORY 1: PUBLIC FINANCE
AUGUST 2013

CESifo *Working Paper No. 4359*

The Intergenerational Welfare State

Abstract

The welfare state is not merely a stand-in for missing markets; it can do a whole lot more. When generations overlap and the young must borrow to make educational investments, a dynamically-efficient welfare state, by taxing the middle-aged and offering a compensatory old-age pension, can generate higher long-run human capital and welfare compared to laissez faire. Along the transition, no generation is hurt and some are better off. If an intergenerational human capital externality is present, unfunded pensions can be gradually phased out entirely. Public pension reform can be rationalized on efficiency grounds without relying on political-economy concerns or aging.

JEL-Code: H000.

Torben M. Andersen
Department of Economics
University of Aarhus
Denmark – 8000 Aarhus C
tandersen@econb.au.dk

*Joydeep Bhattacharya**
Department of Economics
Iowa State University
USA – Ames IA 50011-1070
joydeep@iastate.edu

*corresponding author

Revised July 31, 2013
The authors thank Ron Lee and Tim Miller for sharing their data and acknowledge general financial support from the Danish Council for Independent Research (Social Sciences) under the Danish Ministry of Science, Technology and Innovation. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

In a welfare state, the public sector makes significant tax-financed spending on education, health and old-age pensions. Every developed country is a welfare state, and most have been so for over a century. The United States is no exception; here institutions of public education were established well before those of public pensions, and to this day, among all public transfer programs, education-related ones are the largest.[1] And yet, the term 'welfare state' mainly evokes negative associations associated with entitlement programs, primarily social security[2]; public education and its success story is largely forgotten. Indeed, public education is "generally seen as an investment in human capital, but rarely as an intergenerational transfer" (Bommier et. al., 2010).

A casualty of the presumed synonymity between the welfare state and social security is that any discussion of the life-cycle pattern of intergenerational transfers remains restricted to the middle and old ages, even though education-related transfers, by virtue of their timing in the life cycle, are far more important than old-age transfers, such as social security.[3] The fact of the matter is, in a welfare state, tax-financed welfare programs exhibit a clear age profile over the *entire* lifecyle: both the young and the old are typical net beneficiaries, while the middle-aged are net contributors. This sort of age dependency naturally arises because the need for expenses on education, health and pensions is firmly dependent on age, and because most revenue is raised via taxes on the working, middle-aged. Figure 1 illustrates this property. It shows age-dependent net contributions (taxes minus transfers) for a group of twenty countries, the U.S. and the Scandinavian countries; these exhibit a pattern: net benefits to the individual are positive as young and old, and negative in the intermediate years.

[1] Bommier et. al (2010) report that, for 2008, education (including higher education) was the largest public transfer program in the U.S., at 5.2% of GNP, with Social Security and Medicare at 3.5% and 3.2% respectively.

[2] See Brooks (2012) and the references cited in Folbre (2012). This quote from Samuelson (2011) exemplifies the assumed equivalence between public pensions and the welfare state: "The modern welfare state has reached a historic reckoning. As a political institution, it hasn't adapted to change. Politics and economics are at loggerheads. Vast populations in Europe and America expect promised benefits and, understandably, resent any hint that they will be cut."

[3] Bommier et. al (2010) argue that even after "taking survival probabilities into account, to a recipient a dollar of educational benefits can easily be worth 10 dollars of old-age benefits."

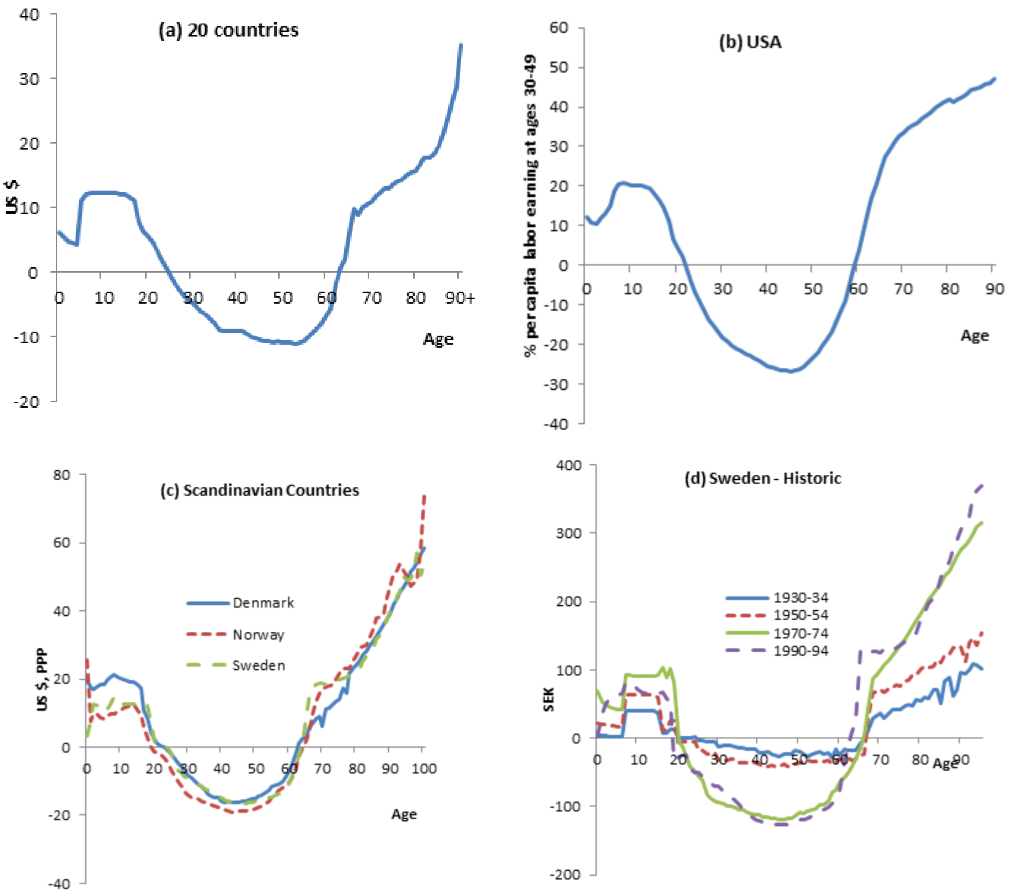

Figure 1: Age-dependent net transfers between the individual and the welfare state

Notes on Figure 1: (a) 20 countries around 2000. Net Public transfers are defined as inflows including services in kinds and cash transfers less outflows including all forms of taxation and social security contributions. Units of 1,000 US $. Source: Miller (2011). (b) Data applies to 2003, and gives the sum of net public transfers and public asset-based reallocations. Units of 1,000 US $. Source: Lee and Mason (2011b). (c) Data for Denmark and Norway applies to 2009, and for Sweden 2008. Swedish data is adjusted for wage increases in 2009, and data is presented in US $ using OECD PPP exchange rates. Units of 1,000 US $. Source: Danish Economic Council (2012), Norwegian Ministry of Finance, Swedish Ministry of Finance (2011) and www.oecd.org. (d) Calculated as averages over 5 years intervals, i.e. 5-10 is the age groups above 5 years and below 10 years. Source: Petterson et.al. (2006).

Figure 1d shows how, for Sweden, the nature of the lifecycle profile of net contributions has changed *over time* – from low early-on benefits and modest pensions to high values of both, funded by increasing net contributions from the middle-aged. The overall picture is clear: the net transfer from the welfare state to the individual is clearly positive early on, and again late in life, and is clearly negative in the working years.

An associated question of great practical significance is: under current welfare arrangements, is the lifetime expected present value of *net* welfare receipts positive for an individual? That is, does the expected present value of publicly-funded care, education and health benefits received early in life dominate the cost of subsequent tax payments as well as care, health and pensions received later in life?[4] For Denmark – see Danish Economic Council (2012) – even respecting fiscal sustainability, the expected PV to the average Dane is about \$80,000, meaning the gain to an individual from participating in the Danish welfare arrangement (measured in net present value terms) is positive. The implication is stark: under fiscal sustainability, the Danish welfare state adds to lifetime income (expands the budget set) of an average Dane by about \$80,000, roughly 5% of the average present value of lifetime earnings. A similar picture – see Figure 2 – emerges for the United States. Bommier at. al (2010), from whom Figure 2 is borrowed, calculate the net present value for each transfer program and each birth cohort as the difference between the lifetime-discounted, survival-weighted benefits and the lifetime-discounted, survival-weighted tax payments for these programs. Figure 2 presents net present values for education and Social Security cum Medicare as a percent of the present value of lifetime earnings.[5]For much of the post WWII era, the *combined* net present value

[4]Denote by n_a the net benefit (taxes minus transfers) at age a, and let the fraction of a cohort alive at age a be π_a. From a life-time perspective of a cohort, the present value of the net benefits from the scheme is given by $PV \equiv \sum_a \left(\frac{1}{1+r}\right)^a \pi_a n_a$, where r is the discount rate. Impose further, a condition ensuring fiscal sustainability: for a stationary population, a pay-as-you-go (PAYG) scheme is fiscally sustainable if $\sum_a \pi_a n_a = 0$. When a welfare arrangement is *front-loaded*, as Figures 1a-d seem to indicate, $PV > 0$. A more detailed discussion is contained in Appendix A.

[5]Bommier et. al (2010) argue that the "creation of Social Security in the late 1930s (with regular benefit payments starting in 1950) and of Medicare in the mid-1960s led to large windfall gains for the early participants in these pay-as-you-go systems. These early participants received benefits far in excess of the taxes they paid for these programs." Why the negative numbers for education? Again, as Bommier et. al (2010) explain "as enrollments and median grade attainments rose, each generation of taxpayers funded a higher level of education than it received itself, so net present values were negative. The generations that funded the education of the baby boom cohorts were heavily taxed because there were so many students and relatively few taxpayers and because enrollment increases were particularly rapid. Those generations born between 1928 and 1942 experienced losses of at least 5 percent of lifetime earnings through the transfer effected by the educational system."

of expenditures on education and Social Security plus Medicare has been in the positive territory.[6]

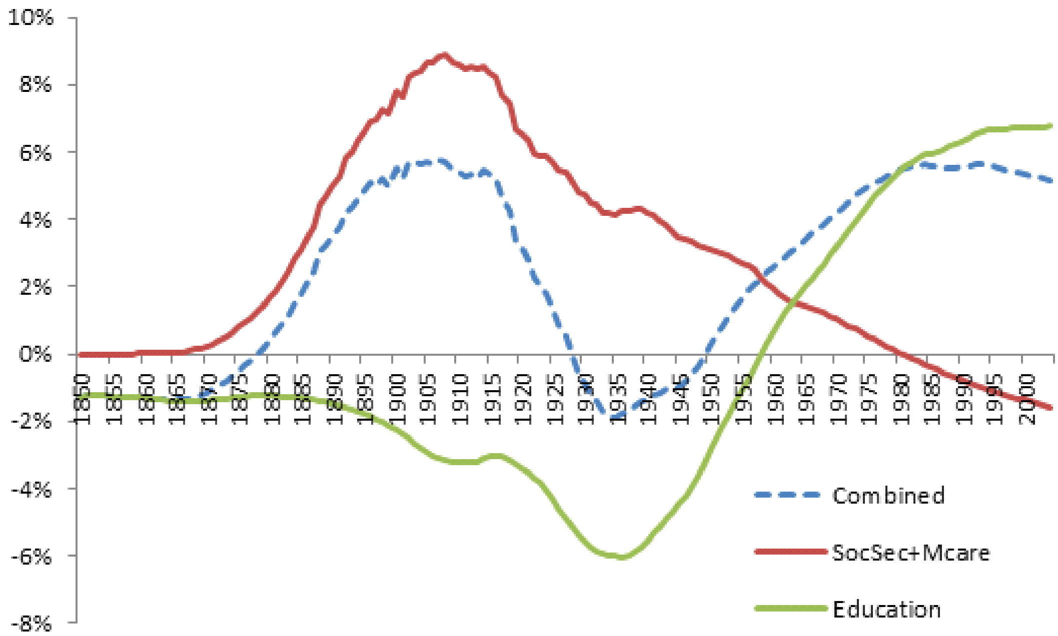

Figure 2: Expected present value of the net-transfers between the individual and the welfare state, U.S. 1850-2004

Notes on Figure 2: the expected present value of public net transfer over the life-cyle as a percent of expected life time income. The inflation-adjusted interest rate is 3% and productivity growth is 1.6%. Source: Bommier et. al (2011).

The assumed interchangeability of public pension programs with the welfare state has an academic dimension as well. For if the welfare state is little more than pay-as-you-go (PAYG) social security, then the famed Aaron-Samuelson result in dynamic public finance offers a clear indictment of the welfare state, at least from an efficiency standpoint. The Aaron-Samuelson argument (see Aaron, 1966 or Feldstein and Leibman, 2002) is that introduction of such a pension scheme in a

[6]See Figure 7 in Bommier et.al (2010). It is worthwhile to note that "the first generations to bear the cost of public education were too old to gain from the introduction of Social Security." Matters were different for later generations. For example, for the cohort born in 1926, "net Social Security and Medicare benefits amounted to 5.5 percent of lifetime earnings, which were offset by a net public education benefit amounting to –4.6 percent of lifetime earnings, so that the net effect of all transfer systems was just +0.9 percent of lifetime earnings."

dynamically-efficient economy, the empirically relevant case, benefits the inaugural generation at the detriment of all future generations. As such, if such a PAYG scheme is in place, then presumably it's best to either leave it be or phase it out and (possibly) replace it with a fully-funded pension scheme. In this context, an aspect of Figure 2 worth highlighting is that, for the United States, net Social Security and Medicare benefits as a fraction of lifetime earnings has been steadily declining – from a peak of around 8% (for the cohort born in 1914) – for the past eighty years. A similar story can also be read from Figure 1d; notice, how in recent years, the net contribution by the retired in Sweden is becoming *less negative* over time. Indeed, world over, the public pension system is under attack. Almost everywhere, across countries with differing political structures and old-age dependency ratios, sharp reductions in public pension promises are being planned. In many cases, transition to a system where public pensions are complemented with (or even replaced by) fully-funded arrangements is underway.[7],[8]

The upshot of the preceding discussion is the following set of 'stylized facts' about welfare states. Any modern welfare state has two arms.[9] The aforediscussed age-dependency of net contributions suggests a *two-part* intergenerational social contract is operative – the middle-aged (the parents in any generation) support welfare payments, both for the young (their children) and the old (their parents). Welfare arrangements are typically *front-loaded* in the sense that the present value of net welfare receipts is positive for the individual. This suggests the "borrowing" component of the intergenerational arrangement – borrowing from the middle aged to finance education for the young – is considerably more important than the "saving"

[7]The OECD Pensions Outlook 2012 argues that pension reforms since the mid 1980s have led "to a reduction in public pension promises in many (OECD) countries, typically between a fifth and a quarter", that these cuts "call for longer working periods and an expanded role for funded, private pensions", and that most OECD countries "have already moved or are moving towards a more diversified system, where PAYG pensions need to be complemented with fully funded pension arrangements..." Arguably, what has precipitated these calls for pension reform in public discourse is the aging of populations in advanced economies. As we demonstrate below, there may be solid efficiency reasons, unconnected with aging, that may yet justify a phasing out of PAYG pensions.

[8]Denmark, the first country to introduce public pensions more than a century back, and by most counts a classic welfare state, made a move in this direction in the late 1980s introducing a fully-funded mandatory labor market pension scheme. Sweden made a similar move in the early 1990s.

[9]Garfinkel and Smeeding (2010) define welfare states as countries with prominent social institutions designed to reduce the inevitable "economic insecurity" produced by the market economy. By this definition, they argue "public expenditures on education, health, insurance, and cash benefits (social insurance and public assistance) all reduce economic insecurity, and hence, ought to considered integral parts of the welfare state." Lindert (2004) also makes a strong case for using this sort of definition. Barr (2001) says "... the term 'welfare state' is used for the state's activities in three broad areas: income transfers, health and health care, and education."

element – sponsoring a pension for the old so as to receive one in the future from the then middle-aged. Finally, by most counts, the public pension component of welfare states with very different political structures and aging patterns is rapidly dwindling. A key challenge for economists is to stay cognizant of these stylized facts and yet explain this nearly-ubiquitous pattern of evolution of the welfare state.

The objective of this paper is to take on this challenge and develop a theory of the intergenerational welfare state that can also rationalize pension reform on pure efficiency grounds, without relying on political-economy concerns or the aging of populations.[10] Unlike previous work characterizing welfare states as stand-ins for missing markets or as instruments of 'market repair', the paper shows, by exploiting front-loadedness, the "borrowing" component of the intergenerational arrangement, a dynamically-efficient welfare state can *improve* over complete market outcomes. To do so in a manner that does not hurt any generation, it must offer, as compensation, the saving component, a pension to the old. If by exploiting the front-loadedness, the welfare state succeeds in unleashing some welfare gains early on – a tail wind – it can even allow one of its arms, the pension payments to the old, to wither away. The central thesis is, while public education needs to be supported and expanded, there may be solid efficiency reasons that justify a gradual phasing out of public pensions.

A sketch of the specifics is in order. We study a three-period overlapping-generations model of a small open economy populated by self-interested agents, similar in many respects to the one studied by Boldrin and Montes (2005), hereafter "BM". In the present model, the young can access the international capital market at fixed, gross interest $R > 1$ (implying dynamic efficiency; no population growth) to finance their investment in human capital. When middle aged, they pay off past loans, work at a competitive wage, consume, and save in that same market at return R. When old, they consume their wealth and perish. There is a single good produced using a standard, neoclassical technology using physical and human capital as inputs. Human capital of the young depends on current education-related spending (both private and public); in some formulations, it depends as well on the human capital of their parents – a positive intergenerational externality. For now, publicly-funded education and private expenses on education are assumed to be perfect substitutes in the production of human capital.

[10] See Razin, Sadka and Swagel (2002) for a model connecting increases in dependency ratios with public pension reforms.

Consider a government that contemplates offering free public education for the young (as top-up on their own spending) financed by a lump-sum tax levied on middle-aged incomes. We show, in a steady state, even under idealized conditions (complete loan markets and perfect substitutability of public and private education expenses) there is a welfare rationale for front-loaded PAYG education for the young.[11] (Parenthetically, the Aaron-Samuelson result argues there is no welfare raison d'etre for PAYG pensions under the same conditions.) The logic here is simple: under dynamic efficiency because $R > 1$, it is cheaper to "borrow" from the government (with associated tax obligations at opportunity cost of 1) than to borrow from the market (with associated interest obligations at opportunity cost of R). Indeed, under perfect substitutability of public and private education expenses, and with $R > 1$, optimal government intervention in education necessitates driving private education spending to the zero corner. In this case, aggregate human capital with 100% public education is at least as high as what is chosen with 100% private education under perfect capital markets because of the interest savings (R versus 1) on education loans. Even if the human capital externality is absent, the steady-state welfare of a representative, two-period lived generation – the Golden rule level – is higher than what is achievable under complete private markets when *all* education spending is mediated by the government.[12,13] Clearly, a *range* of public-education spending levels exist, from a lower bound of that achievable under complete private markets to the upper bound level of the Golden rule; assume here on, the government fixes forever a public-education spending in the interior of this range – call that level of spending, G.

Right at the onset, the classic implementation problem makes its appearance. How to rally support for G from the inaugural middle-aged generation, them never having received similar public assistance in the past. Inspired by Becker and Murphy (1988) and BM, we propose the following implementation strategy: each period, collect taxes from the middle-aged to finance G for the current young, and in the following period, tax the by-then middle-aged to finance, both a lump-sum pension (call it P) to the by-then elderly *and* G for the newly born. Within such a intergener-

[11] A more detailed discussion is contained in Appendix A.

[12] If the human capital externality is present, the level of education (and welfare) achievable by a benevolent government, one that maximizes steady-state welfare of a representative, two-period lived generation, is, of course, even higher.

[13] In 2011, an average of 91% of primary, secondary and post-secondary non-tertiary education in OECD countries, and never less than 80%, is paid for publicly. Moreover, the trend in these expenditures is mildly upward. See OECD (2011).

ational compact, each generation invests in the human capital of the next generation and is itself taken care of at the end of its life by the generations in which it has invested. Specifically, the government becomes a benevolent welfare state with two arms, public education and public pensions. To study whether this implementation strategy works, following Kotlikoff (1998) and many others, we adopt the Pareto criterion as a gate-keeper for implementability. That is, we ask, can the government introduce a education-pension package – G and P – and still ensure all agents alive at the onset, and those to come, experience at least the same level of utility under the package as they would in its absence?

We show, for a range of G, an appropriate pension compensation – one that makes up, at least, the opportunity cost of taxes paid (and forgone return on saving) – can be constructed under the daunting Pareto criterion. By promising such a path of pensions to the inaugural (and subsequent) middle-aged generation(s) in their old age, it is possible to get such a education-pension (henceforth, "EP") package off the ground. And, it is the ongoing promise of a compensatory old-age pension that helps sustain support for public education. The challenge then becomes, how to get our EP package, one that seeks to implement more education spending than under complete markets, off the ground *and* have the pension component wane, all under the Pareto criterion.

For this to work, somehow, the support for G, a welfare-enhancing activity, must sustain itself, needing less and less help from compensatory pensions. This is where the aforediscussed intergenerational human capital externality kicks in to generate a "tail wind". Compared to the complete markets outcome, the higher level of education spending under the public scheme produces a higher level of human capital for the current young and, by the same stroke, raises the human capital of the next round of parents, and so on for subsequent generations. The higher human capital levels translate into higher wage incomes with accompanying potential for welfare gains. We show, more strikingly, the path of P is non-monotonic – a period of initial rise followed by a sustained fall. And it is the welfare gain from education fueled by the human capital externality – the tail wind – that helps bring down P.[14] The message is clear: while public pensions play an important role in getting the EP package off the ground and solving subsequent implementation problems, they don't have to stick around. From a forward-looking perspective, they can be

[14] The increased and continuing need to invest in the human capital of future generations and its connection with the pension reform debate is a major theme in Bovenberg (2010).

phased out eventually, at no welfare cost, and replaced, possibly, with fully-funded pensions.

The plan for the rest of the paper is as follows. In Section 2, we start by reviewing the literature taking care to compare our work with BM. In Section 3, we set up the model economy, establish the complete market allocations, and derive the Golden rule allocations. In Section 4, we study EP packages in a steady state, and in Section 5, the associated implementation hurdles. Section 6 concludes. Proofs of major results as well as some extra clarifying material and additional results are contained in the appendices.

2 A review of the literature

The idea of the desirability of a welfare state with a two-part social contract is not new and certainly goes back, at least, to Pogue and Sgontz (1977) and Hammond (1975).[15] Becker and Murphy (1988) clarify these ideas further and set the stage for the implementation issue in Boldrin and Montes (2005). They argue that altruistic parents may not invest efficient amounts in their children if those investments compete with their own retirement provisions. The welfare state could correct this deficiency by taxing parents and financing public education. The implementation problem arises because the parents are now worse off than before, having been forced to spend more on their children's education than they wished. To compensate them, the welfare state taxes the children later in life to pay their by-now old parents a public pension.[16]

Another important contribution is Rangel (2003) who addresses the issue of support for a policy package via trigger strategies. He allows for backward (to the old) and forward intergenerational goods/exchanges (to the young) in PAYG-

[15] Political scientists, prominently Esping-Andersen (1990), have argued that advanced industrial societies have produced three kinds of welfare states – liberal, conservative, and social democratic. In the liberal cluster, solutions are market based and welfare coverage is minimal and means-tested. The conservative nations support preservation of the status quo; social protection is tied to labour market status. In the social democratic cluster, welfare coverage is universal and every attempt is made to guarantee the individual and families "a socially acceptable standard of living independently of market participation" without discouraging individual aspirations or denying them opportunities for advancement. Such classifications, while no doubt useful, tend to ignore the intergenerational nature of the welfare state, our primary focus.

[16] Both Konrad (1995) and Poutvaara (2003) push the Becker and Murphy (1988) ideas forward and argue that even in the absence of altruism, the old may support spending on the young if such action indirectly benefits the old via general-equilibrium changes in asset prices or the tax base. In the current paper, we hold the interest rate fixed thereby eliminating any general-equilibrium effects via that route. General equilibrium effects on the interest rate are of paramount importance in establishing political support for social security in Gonzalez-Eiras and Niepelt (2008).

financed schemes. He argues it is possible to support an EP package under a trigger strategy – a necessary condition is that the continuation surplus of the backward part is positive. This ensures that the middle-aged do not have an incentive to break the contract – their financing of education is ensured by the threat of losing the pension! Rangel (2003), however, does not include general equilibrium effects in the analysis. Nor does he consider the implementation problem.[17]

It is most instructive to compare our results with those in BM, the paper nearest in spirit to the current one. In their important and influential paper, BM compare two settings, one in which education-loan markets are complete with one where they are totally absent.[18] In the latter setting, they consider the following EP package: tax the middle aged to finance the same level of education as is privately optimal

[17] Our use of the Pareto criterion, establishes, in a sense, the set of policies that are least likely to face stiff resistance as they make their way through standard democratic processes. There is, by now, a large body of work in the political-economy domain that studies the co-existence of the twin institutions of public education and public pensions. Cooley and Soares (1999), Boldrin and Rustichini (2000), Conde Ruiz and Galasso (2003), Poutvaara (2006), and others study conditions under which public education and social security can be jointly supported with repeated voting by a social contract relying on subgame-perfect voting strategies ("trigger strategies") without commitment. In an interesting paper, Hassler et.al (2004) show that political support for distortionary redistribution can persist when agents make irreversible human capital decisions which affects their subsequent political preferences. In their setting, if agents believe that the breakdown of the welfare state is possible, they vote strategically so as to induce, in the future, a change in the identity of the median voter and bring about the demise of the welfare state.

Kaganovich and Zilcha (2012) analyze how the support for tax-financed education depends on the nature of the pension scheme (fully funded vs. PAYG). A higher level of education increases the return to capital as well as the tax base. The political equilibrium has higher education under a fully-funded scheme due to the positive rate-of-return effect. In a PAYG scheme, by contrast, the larger tax base is countered by the higher rate of return thereby making the pension scheme less attractive.

Bosi and Gumus (2012) study the political support for simultaneous intergenerational transfers in the form of pensions to the old and intragenerational transfer payments to those with low incomes. Iturbe-Ormaetxe and Valera (2012) do not study the co-existence of public education and pensions. Instead, they focus on the labor market consequences of pension reform and how these affect political support for public education. Lancia and Russo (2012) push the Rangel (2003) line of thinking forward by including backward and forward intergenerational transfers – pension benefits and higher education investments – in a repeated-voting setup of electoral competition with endogenous human capital accumulation and labor income taxation. They assume that private capital markets supplying education loans do not exist, but, like us, they assume an education externality. They, compare the politico-equilibrium outcome to that chosen by a benvolent government. As in Docquier et. al (2007), the discounting of different cohorts matters. They do not characterize the Pareto-improving policies as we do. Interestingly, they find that dynamic efficiency is necessary for both forward and backward intergenerational transfers to simultaneously arise. Naito (2012) studies a stationary Markov-perfect equilibrium in which social security and public education is jointly determined and complementarity between public education and overall human capital accumulation is key.

[18] It has been noted that their assumption of complete absence of education-loan markets for the young is somewhat extreme. Andalfatto and Gervais (2006), and more recently, Wang (2011) ask if the Boldrin-Montes EP package replicates the complete market allocation in the presence of *endogenous* borrowing constraints. The effects, as Wang (2011) points out, are not obvious. In such models, a generous EP package dims the desire for consumption smoothing between middle and old age. This increases the incentive to default on past loans, and results in anticipated stricter borrowing limits on the young.

under complete markets and pay the old a pension that compensates them at market interest rates for their prior tax sacrifice. BM show that such an EP package can exactly replicate complete market allocations in a world in which education-loan markets are missing. The implication is, if complete education-loan markets are present, there is no efficiency rationale for a welfare state.[19] In a sense, we argue the welfare state can do *much more*! Even when private credit markets are complete, the state can intermediate the lending activity at the "biological interest rate" which is lower than what the market charges, thereby generating an efficiency gain. This "borrowing argument" allows the state to not only replicate the level of human capital attainable under complete markets but surpass it.

The Boldrin-Montes EP package, like ours, confronts the implementation problem head on. In their case though, the pension compensation needed to float the education scheme and leave every generation indifferent to laissez faire, cannot dwindle over time. This is true even if a human capital externality is present. For us, the pension compensation can start to decline because of the higher human capital generated by public education. It is also noteworthy that under the Boldrin-Montes EP package, the present value of net welfare receipts is zero for the individual implying that the aforediscussed "borrowing" component of the intergenerational arrangement exactly matches the "saving" element. In contrast, our EP package is shown to be front-loaded, and hence, is consistent with current welfare arrangements (see Figure 1) in most countries.

Docquier, Paddison, and Pestieau (2007) and Del Rey and Lopez-Garcia (2010), in slightly different model environments, compare physical and human capital levels in a complete competitive economy with those attainable by a social planner maximizing a discounted sum of individual utilities for an arbitrary social discount rate. Their primary goal is to study whether problems relating to over or under accumulation of physical and human capital can be corrected via appropriate education subsidies or old-age pensions. In direct contrast to results in Docquier, Paddison,

[19] It is important to recognize that Boldrin and Montes (2005) study a closed economy. Therein they show that, in the absence of education-loan credit, the twin public instituitions of education and pensions can replicate the complete markets allocation. They claim, though, under reasonable parameter specifications, the complete markets equilibrium is dynamically efficient because the marginal product of capital exceeds the economy's growth rate. Docquier, Paddison, and Pestieau (2007), Del Rey and Lopez-Garcia (2010) and Bishnu (2011) take issue with that characterization. Their basic argument is, whether the rate of return to capital is higher or lower than the economy's growth rate is not sufficient to assess dynamic efficiency; even when the economy accumulates too little capital, it may overinvest in human capital. Docquier, Paddison, and Pestieau (2007) also point out that Boldrin and Montes use a "restrictive" notion of efficiency, one that ignores the spillover effects from the human capital externality.

and Pestieau (2007), Del Rey and Lopez-Garcia (2010) find that simultaneous underaccumulation of physical capital and overaccumulation of human capital is possible. Moreover, the optimal education subsidy is negative, and in some settings, a case can be made for pensions to the old. Docquier, Paddison, and Pestieau (2007) categorically reject any efficiency-enhancing role for pensions.[20]It bears emphasis that these papers do not focus on the implementation problem and do not impose the Pareto criterion generation by generation.

Our work has some bearing on the literature studying the transition from a existing PAYG system to a fully funded one, and whether a Pareto-improving transition to a fully funded system is possible. In that literature – see Lindbeck and Persson (2003) and references therein – it is typically assumed that the PAYG system, for whatever historical reason, is already in place, and because of the aforediscussed Aaron-Samuelson logic, it ought to be phased out. For us, the PAYG pension scheme is needed for a reason (as compensation for prior contributions to the funding of public education). We argue that having served its purpose, by unleashing enough of a tail wind, it can be phased out, and that too in a Pareto-neutral manner. Of course, there may be many good reasons – distributional, intergenerational risk-sharing, old-age poverty reduction, and so on – that may justify the continuance of the PAYG pensions. Our argument for the phasing out of public pensions is purely related to dynamic efficiency concerns.

3 The model

3.1 Primitives

We consider a one-good, small open economy consisting of an infinite sequence of three period-lived overlapping generations, and a government. At each date $t = 1, 2, .., \infty$, a continuum of agents with unit mass is born (no population growth). When young, agents can access complete private markets to secure loans (D) that finance acquisition of human capital, an asset which provides no direct utility but affects future earnings. In their middle age, they pay off these loans, work at a

[20]In an interesting twist to this literature, Bishnu (2011) finds that human capital externalities are superfluous in the sense that consumption externalities alone may provide the necessary rationale for government intervention in education. He too finds that both under and over accumulation of physical and human capital is possible with intergenerational consumption externalities. However, the possibility that the laissez-faire equilibrium may deviate from the social optimum in its allocation of physical and human capital, in opposite directions, is no longer present. In his setup, whenever a education subsidy is justified, a PAYG pension is not.

competitive wage, w, consume, and save for old age in the capital market at a fixed return, $R \geq 1$. (The borrowing rate, in this case, the rate of interest on education loans, is also R.) When old, they consume their entire wealth and die, leaving no bequests.[21]

Let c_t^m (c_{t+1}^o) denote consumption of the final good by a representative middle-aged (old) agent born at $t-1$. All such agents have preferences representable by the utility function $u(c_t^m) + \delta\, u(c_{t+1}^o)$ where u is twice-continuously differentiable, strictly increasing, and strictly concave in its arguments, the usual Inada conditions hold, and $\delta \in (0,1)$.[22]

The government is a benevolent welfare state with (potentially) two arms: it offers a) free education, $G \geq 0$, for the young (as a top-up on their own education spending, D) and b) a pension $P \geq 0$ for the old. Each generation contributes twice – for the previous and the succeeding generation – and benefits twice, as a child and in old age. This EP package is financed by a lump-sum tax, T, levied on the income of the middle-aged.[23] To disentangle the effects of pensions and education expenses, the lump-sum tax is divided into two parts, the part financing pensions and the one financing education:

$$T_t = T_t^G + T_t^P = G_t + P_t;\ T_t^G = G_t;\ T_t^P = P_t. \tag{1}$$

The seeds for the intuition of many of our results below are sown right here: under a PAYG scheme, government-sponsored education has an opportunity cost of 1 (since the net population growth rate is zero), while the same, when administered by the market, faces an opportunity cost of $R > 1$.

We posit a general form for the process of human capital accumulation:

$$H_t = h\left(E_{t-1}, H_{t-1}\right), \quad E_{t-1} \equiv D_{t-1} + G_{t-1} \tag{2}$$

[21] The absence of altruistic motives is not limiting per se because parents, in any case, cannot internalize the impact of an increase in the aggregate stock of human capital. Kaganovich and Zilcha (1999) study a model with altruistic parents who invest in their offsprings' human capital. Upon retirement, the labor income of the young are taxed to finance a pension to the by-then old parents.

[22] Our results do not rely on additive separability nor on felicity functions at each age being identical.

[23] In keeping with the spirit of BM, government spending does not directly affect production via, say, a Barro (1990)-style production externality. Also, as in BM, the government is not allowed to finance EP packages by borrowing on the international market. Historical experience suggests government borrowing in capital markets was mainly used to finance wars, and never to finance public education schemes.

where H_t is the human capital of a middle-aged agent, E_{t-1} denotes total (private plus public) education expenses on that agent when young, and H_{t-1} is predetermined. Note that private and public education expenses are assumed to be perfect substitutes; what determines human capital is their sum, E, but not its composition. Henceforth, assume

Assumption 1

$$\begin{aligned} h(0,0) &> 0; \frac{\partial H_t}{\partial E_{t-1}} \equiv h_E > 0, \frac{\partial^2 H_t}{\partial E_{t-1}^2} \equiv h_{EE} < 0, \frac{\partial H_t}{\partial H_{t-1}} \equiv h_H \in (0,1), \\ \frac{\partial^2 H_t}{\partial H_{t-1}^2} &\equiv h_{HH} < 0, \frac{\partial H_t}{\partial E_{t-1} \partial H_{t-1}} \equiv h_{EH} \geq 0. \end{aligned}$$

The interpretation of Assumption 1 is straightforward. Spending more on education raises the human capital of the young at a declining marginal rate. If $h_H = 0$, no intergenerational human capital externality (HCE) is present. If $h_H > 0$, an intergenerational HCE is present; it affects not just current but all future decisions regarding private education spending. Higher human capital of the previous generation raises the marginal contribution of education spending by the current young.

For future reference, it will be useful to consider a steady-state version of (2), if it exists. To that end, define $H(E) \equiv h(E, H(E))$. The steady-state analog of Assumption 3.1 is

$$H(0) > 0,\ H_E > 0,\ H_{EE} < 0,\ H_{EH} > 0. \tag{3}$$

For reasons that will be evident shortly – see Footnote 30 and Appendix B – we also impose

Assumption 2

$$h_{EE}(E, H(E)) + h_{EH}(E, H(E))\, H_E(E) < 0.$$

3.2 Private decisions

Let S denote saving in middle age. The budget constraints faced by a young agent are given by

$$0 \leq D_{t-1} \leq \frac{wH_t}{R} \tag{4}$$

$$c_t^m = wH_t - T_t - RD_{t-1} - S_t \tag{5}$$

$$c_{t+1}^o = RS_t + P_{t+1};\; P_{t+1} \geq 0 \tag{6}$$

The assumption of complete private markets implies there is no asymmetry between the interest rate paid on borrowing and that received on saving. We proceed under the assumption that the upper-bound constraint on borrowing is non-binding, i.e., $D_{t-1} < \frac{wH_t}{R}$. For future use, note that (5)-(6) and (1) combines into

$$c_t^m + \frac{1}{R}c_{t+1}^o \leq Y_t \equiv wH_t - RD_{t-1} + I_t \tag{7}$$

where $I_t \equiv -G_t - P_t + \frac{1}{R}P_{t+1}$ is the present value of direct financial transactions between the individual and the welfare state.

With pre-determined H_{t-1}, a young agent's problem is given by

$$\max_{S_t, D_{t-1} \geq 0} \quad u\left(wH_t - T_t - RD_{t-1} - S_t\right) + \delta u\left(RS_t + P_{t+1}\right)$$

subject to (2)-(6) and the usual non-negativity constraints on consumption. The saving decision is standard and given by

$$u'\left(c_t^m\right) = R\delta u'(c_{t+1}^o). \tag{8}$$

The only remaining decision-variable of interest is education expenditures. Clearly, for given G_{t-1} and pre-determined H_{t-1}, the choice of private education spending is a one-period decision given by

$$\begin{cases} w\, h_{E_{t-1}}(D_{t-1} + G_{t-1}, H_{t-1}) = R \text{ if } D_{t-1} > 0 \\ w\, h_{E_{t-1}}(D_{t-1} + G_{t-1}, H_{t-1}) < R \text{ if } D_{t-1} = 0 \end{cases}. \tag{9}$$

Note that when $D_{t-1} > 0$, given the perfect substitutability between public and private education spending, there is one-for-one crowding-out, i.e., from (9),

$$w\, h_{EE}(\cdot)\left[1+\frac{dD_{t-1}}{dG_{t-1}}\right]=0 \Leftrightarrow \frac{dD_{t-1}}{dG_{t-1}}=-1. \tag{10}$$

As long as agents are borrowing positive amounts on the private market, they will reduce their own borrowing for education, one-for-one, for every unit increase in government top-up spending. It follows that there exists a G, high enough, such that private education expenses are optimally driven to the zero corner. At that level of G, private loan markets remain open but no trade is conducted.[24]

Evaluated at the optima described by (8)-(9), indirect utility of the middle-aged (given their prior education decision) is given as $V(R, Y_t)$, where Y_t is defined in (7), and $\partial V/\partial Y_t > 0$. This implies utility and Y_t comparisons across EP packages are equivalent, a convenience that exists purely because of our small open economy assumption, and one we exploit below.

3.3 Competitive equilibrium definitions – the CMA

A dynamic competitive equilibrium allocation $\{D_t\}$ with **complete markets**, given a policy package satisfying (1) – henceforth, the **CMA** – is described by

$$\begin{cases} w\, h_{E_{t-1}}(E_{t-1}, H_{t-1}(\cdot)) = R \text{ if } D_{t-1} > 0 \\ w\, h_{E_{t-1}}(E_{t-1}, H_{t-1}(\cdot)) < R \text{ if } D_{t-1} = 0 \end{cases}, \tag{11}$$

where $E_{t-1} \equiv D_{t-1}$ in the absence of government intervention and $E_{t-1} \equiv D_{t-1} + G_{t-1}$ in its presence; additionally (2)-(6) hold. Specifically, in the absence of any government policy intervention, and with an unfettered education-loans market, private agents choose a level of education consistent with

$$\begin{cases} w\, h_{E_{t-1}}(\tilde{D}_{t-1}, H_{t-1}(\cdot)) = R \text{ if } \tilde{D}_{t-1} > 0 \\ w\, h_{E_{t-1}}(\tilde{D}_{t-1}, H_{t-1}(\cdot)) < R \text{ if } \tilde{D}_{t-1} = 0 \end{cases}. \tag{12}$$

Henceforth, we define $\tilde{D}_{t-1}$ as the CMA level of private education spending in the absence of government intervention.

Apropos the discussion surrounding (10) above, it follows that if G rises, D falls, and hence private and public education continue to coexist iff $0 < G_{t-1} < \tilde{D}_{t-1}$

[24] This is distinct from BM's notion of perfectly incomplete markets; there, private markets do not exist and, naturally, there is no trade.

where (12) defines $\tilde{D}_{t-1}$; then $E_{t-1} \equiv G_{t-1}$ and $D_{t-1} = 0$, in which case

$$w\, h_{E_{t-1}}(G_{t-1}, H_{t-1}(\cdot)) < R \tag{13}$$

holds.

Noting from (2) that $H_t = h(E_{t-1}, H_{t-1})$, the steady-state analog of (12), henceforth the **CMA steady state** – may be written as

$$\begin{cases} w\, h_E(\tilde{D}, H(\tilde{D})) = R \text{ if } \tilde{D} > 0 \\ w\, h_E(\tilde{D}, H(\tilde{D})) < R \text{ if } \tilde{D} = 0 \end{cases}, \tag{14}$$

where $H(\tilde{D}) \equiv h(\tilde{D}, H(\tilde{D}))$. Again, if $0 < G < \tilde{D}$, private and public education continue to coexist in the long run and $w\, h_E(E, H(E)) = R$. If $G \geq \tilde{D}$, private education is at a corner; in that case, $E = G$ with $wh_E(E, H(E)) < R$.

3.4 Optimal education-pension package in steady state: Golden rule

Recall, maximizing lifetime welfare is equivalent to maximizing Y (see (7)). In steady state, Y reduces to

$$\Delta(G, P, D) \equiv w\, h(D + G, H(D + G)) - RD - G + \left(\frac{1}{R} - 1\right) P.$$

The Golden-rule optimality program – a EP package that maximizes steady-state lifetime welfare of any representative three-period lived generation – is described by $\max\limits_{G \geq 0, P \geq 0} \Delta(G, P, D)$. The following lemma outlines Golden rule policies.

Proposition 1 *Any Golden rule EP package has $E = G^{gr}$ and $D = 0$. In the absence of human capital externalities, the Golden rule level of G is defined by $wh_E(G^{gr}) \equiv 1$. In the presence of human capital externalities*[25]*, the Golden rule*

[25] Since $H = h(D + G, H)$, it follows that

$$\frac{dH}{dG} = h_E(E, H) \frac{dE}{dG} + h_H \frac{dH}{dG}$$

implying

$$\frac{dH}{dG} = \frac{h_E(E, H)\left(1 + \frac{dD}{dG}\right)}{1 - h_H}.$$

level of G is defined by

$$w\left[\frac{h_E\left(G^{gr},H\left(G^{gr}\right)\right)}{1-h_H\left(\cdot\right)}\right]\equiv 1. \tag{15}$$

In either case, the Golden rule level of P is 0.

A few remarks are in order. First, the result that the Golden rule involves no pensions is really a restatement of the classic Aaron-Samuelson result (see Aaron, 1966) in dynamic pension economics. It follows from the idea that unfunded pensions are an inefficient saving instrument when $R > 1$. Second, since P is *constrained* to be non-negative, the Golden rule pension level is 0 and not lower.[26] Third, at the Golden rule, private education spending is necessarily absent. Finally, the presence of the term involving h_H in (15) reflects that the government is taking the HCE into account (something private agents are not). This generates a push for higher accumulation of human capital under the Golden rule than in the CMA. However, the HCE is not necessary to generate this. Even when $h_H\left(\cdot\right) = 0$, $wh_E\left(G^{gr}\right) = 1$ while $wh_E(\tilde{D}) = R$ in the CMA; since $R > 1$, this implies $G^{gr} > \tilde{D}$ or that overall human capital is higher under the Golden rule than in the CMA (purely from the implied interest differential).

4 Education-pension packages: Improving upon the CMA in steady state

In this section, we study publicly-funded EP packages that attempt to improve upon the CMA allocation. Of course, one such EP package would be the Golden Rule package outlined in Proposition 1. But are there others, possibly less demanding? We start by establishing that publicly-funded EP packages – henceforth (G, P) packages – exist that produce such an improvement. That is, via an appropriate (G, P) package, possibly far away from the Golden Rule package, it is possible to improve welfare locally near the CMA steady-state in (14). This naturally leads to the issue of implementation: can such packages be introduced "voluntarily"? As discussed in the introduction, our yardstick for improvement will be the Pareto criterion. Specifically we ask, starting from a steady-state CMA allocation, can the government introduce a (G, P) package that leaves every current and future generation at least

[26] Docquier, Paddison, and Pestieau (2007) advocate for negative pensions on the elderly.

as well off as what they would have been under the CMA?

Suppose the initial setting, *the basis for comparison*, is the CMA steady-state – (14) with no taxes or pensions, $\tilde{D} > 0$ and $w\,h_E(\tilde{D}, H(\tilde{D})) = R$. With private education spending held fixed at $\tilde{D}$, suppose a permanent but small (G, P) package is introduced, and the economy settles down to a new steady state with $E = \tilde{D} + G$, and $H = h(\tilde{D} + G, H)$. We start by asking, is it possible to improve steady-state welfare over the CMA level by changing the (G, P) package? Our criterion for improvement, naturally, is the steady-state lifetime welfare of any representative three-period lived generation, given by

$$\Delta\,(G, P, D) = w\,h\,(D + G, H) - RD - G - P + \frac{1}{R}P.$$

Holding D fixed at $\tilde{D}$, we have

$$\frac{\partial \Delta\,(G, P, D)}{\partial P}|_{D=\tilde{D}} = -1 + \frac{1}{R} < 0.$$

Hence, introducing a pension is a bad welfare move; in fact, since the marginal condition above depends only on R and not on any endogenous variable (see below), introducing a *permanent* pension at any steady state level of education is welfare reducing. And this is true irrespective of whether the HCE is present. Similarly, for fixed P, holding D fixed at $\tilde{D}$, and using $w\,h_E(E, H) = R$, we have

$$\frac{\partial \Delta\,(G, P, D)}{\partial G}|_{D=\tilde{D}} = \frac{R}{1 - h_H} - 1 > 0,$$

implying that, ceteris paribus, a marginal topping up on the CMA education level, $\tilde{D}$, is welfare improving (holds for $h_H = 0$). These local results jointly suggest that, at the CMA steady state, introducing a pension is a bad welfare move while adding to existing education spending is not. Again, these ideas do not rely on the presence of the HCE.[27]

[27] Since $H = h\,(E, H)$, $\frac{dH}{dP} = h_E\,(E, H)\,\frac{dE}{dP} + h_H\frac{dH}{dP}$. With G held fixed, this implies $\frac{dH}{dP} = \frac{h_E(E,H)\left(\frac{dD}{dP}\right)}{1-h_H}$; therefore,

$$\frac{\partial \Delta\,(G, P, D)}{\partial P} = w\frac{h_E\,(E, H)\left(\frac{dD}{dP}\right)}{1 - h_H} - R\frac{dD}{dP} - 1 + \frac{1}{R}.$$

Similarly, for fixed P,

$$\frac{\partial \Delta\,(G, P, D)}{\partial G} = w\frac{h_E\,(E, H)\left(1 + \frac{dD}{dG}\right)}{1 - h_H} - R\frac{dD}{dG} - 1.$$

The intuition for these results is standard. That pensions are not desirable in a dynamically-efficient economy is the well-known Aaron-Samuelson result from public economics. The notion that topping up on the CMA education level is welfare improving follows from, what we term, the "borrowing argument". Recall that the government imposes a lump-sum tax on the middle-aged and uses some of the proceeds to pay for the education expenses of the young. From the perspective of the young, with zero net population growth, \$1 received from the middle-aged today implies a payment (tax) obligation of exactly \$1 to the *then* young next period. Whereas \$1 borrowed on the education-loans market implies a payment of $\$R$ next period. In short, since $R > 1$, it is socially cheaper for education to be funded via intergenerational transfers (at gross return 1) than for it to be funded via the education-loans market (with associated return R).

We have established that the Golden rule package $(G^{gr}, 0)$ improves upon the CMA in steady-state welfare terms. But what about a package $(G, 0)$ with $\tilde{D} < G < G^{gr}$? Do such packages deliver higher welfare than under the CMA?[28] The answer is yes. To see this, note for such a $(G, 0)$ package,

$$\Delta(G, 0, 0) \equiv w\, h(G, H(G)) - G,$$

and hence, $\frac{d\Delta}{dG} = w\frac{h_E(G,H(G))}{1-h_H(\cdot)} - 1$; since $G < G^{gr}$ and $w\frac{h_E(G^{gr},H(G^{gr}))}{1-h_H(\cdot)} - 1 = 0$, it follows that since $\frac{h_E(G,H(G))}{1-h_H(\cdot)}$ is decreasing in G, then $\frac{d\Delta}{dG} = \frac{wh_E(G,H(G))}{1-h_H(\cdot)} - 1 > 0$, implying welfare relative to the CMA is higher at any level of G satisfying $\tilde{D} < G < G^{gr}$. The proposition below summarizes the salient points of the discussion above.

Proposition 2 *A welfare arrangement with $E = G$, $D = 0$, satisfying $\tilde{D} < G < G^{gr}$ and $P = 0$ delivers higher steady-state welfare than what is possible under the CMA.*

To summarize, henceforth we consider packages wherein public education spending at all dates exceeds private spending in the complete markets economy. Additionally, public education spending is set at a level that fully crowds out private education expenses *forever* after. To foreshadow, the main trade-offs that will dominate the discussion below are the following. Down the transition path, the middle-aged will be asked to pay a tax to cover both the public-education expense on the young

[28] Of course, it is possible that $G > G^{gr}$ delivers higher welfare than in the CMA. But that is not our concern here.

and a pension for the old but they will enjoy two benefits, a reduced education-loan interest expense (in opportunity cost terms) and higher human capital that will bring with it higher incomes. But will the inaugural generation buy into this scheme?[29] Specifically, are Pareto improvements over the complete markets outcome possible with public involvement?

5 Implementation

The initial setting, our benchmark for comparison, is the CMA steady state which has no taxes or pensions, $G = 0, D = \tilde{D}$ and $\tilde{H} \equiv H(\tilde{D})$. Consider a EP package, $\Theta_t \equiv \{(G_{t+j}, P_{t+j})\}_{j=0}^{\infty}$ introduced into the CMA steady state at time t where, as discussed above, $G_{t+j} \geq \tilde{D}$ for all j).[30] The issue at hand is, what should Θ_t look like if it is to be implementable? meaning, what sort of a compensation package must Θ_t contain so the middle-aged are not hurt from accepting it?

As stated earlier, our criterion for Θ_t to be acceptable to any generation t is the Pareto criterion:

$$Y_{t+j}(\Theta_t) \geq Y_{t+j}(0) \text{ for all } j.$$

In other words, we ask, given a sequence $\{G_{t+j}\}_{j=0}^{\infty}$, what is the sequence $\{P_{t+j}\}_{j=0}^{\infty}$ $\in \Theta_t$ that leaves every current and future generation at least as well off as what they would have been in the CMA steady state? To keep things simple, we assume the government can costlessly commit to Θ_t.

[29] It is well understood that the earliest participants in the U.S. Social Security system received a windfall gain; similarly, the first generation to finance the public education system incurred a windfall loss because they paid for a level of public educational benefits that they themselves never received. The computations in Bommier et. al (2010) suggest multiple cohorts in U.S. history have experienced net losses through the public transfer systems. They find that those born before 1880 experienced net losses due to the expansion of the public education system. Similarly, those born between 1930 and 1947 also experienced net losses due to the expansion of the pension system. According to their calculations, while these cohorts received large windfall gains associated with the start-up periods for Social Security and Medicare, these were more than offset by windfall losses from the expansion of public education.

[30] Since education-loan markets are complete and private agents adjust their borrowings one-for-one against government top-up spending, an arbitrary $\{G_t\}$ package would, at times correspond to an interior D_t given by $wh_E(D_{t-1} + G_{t-1}, H_{t-1}) = R$, and when G is big enough, to $D_t = 0$ with $wh_E(G_{t-1}, H_{t-1}) < R$. It is readily apparent that this "two-pronged" definition of D would pose serious analytical challenges. To keep things more manageable, choose G to be high enough so that $E = G_{t-1}$ and $D = 0$ at the inaugural date; private education is fully crowded out on impact. The problem is, since the HCE increases the marginal product of human capital investments, further along, a private incentive to undertake education expenses may get reinstated. To prevent this, we require $wh_E(E, H_{t-1}) < R$ for $G > \tilde{D}$ but additionally, we need $wh_E(E, H_{t+k}) < R$ for all $k \geq 0$. Since $H_{t+k} \geq H_{t+k-1}$ for all k, and $h_{EH} > 0$ is assumed, the latter condition may not hold. In Appendix B, we show that Assumption 2 is needed to get this to work.

Middle-age in t: Consider the first generation who are middle aged in period t. They did not receive any top-up on their own young-age education spending, i.e., their education decision is pre-determined. As such, they are at least as well off under Θ_t as in the CMA if

$$wh(\tilde{D}, \tilde{H}) - R\tilde{D} - G_t - P_t + \frac{1}{R}P_{t+1} \geq wh(\tilde{D}, \tilde{H}) - R\tilde{D} \Leftrightarrow I_t \geq I_t^c = 0 \quad (16)$$

where $I_t \equiv -G_t - P_t + \frac{1}{R}P_{t+1}$. (Recall, I_t captures the present value of all financial transactions between a middle-aged (at t) individual and the welfare state.) Of course, $P_t = 0$ since the current old did not contribute to the financing of G and cannot expect a pension. The critical value of I_t which leaves the agent equally well-off as in the CMA steady-state, call it I_t^c, is zero.

Middle-aged in $t+1$: Analogous to above, the no-less utility condition for those who are middle-aged at $t+1$ is given by

$$wh(G_t, \tilde{H}) + I_{t+1} \geq wh(\tilde{D}, \tilde{H}) - R\tilde{D} \Leftrightarrow I_{t+1} \geq I_{t+1}^c \equiv wh(\tilde{D}, \tilde{H}) - wh(G_t, \tilde{H}) - R\tilde{D} < 0. \quad (17)$$

Notice $h(., H_t) = h(G_t, \tilde{H})$, i.e., the stock of human capital is pre-determined at CMA levels. There is a critical value of I_{t+1} which leaves the agent equally well-off as in the CMA steady-state; this critical value, I_{t+1}^c, is negative since i) $wh(G_t, \tilde{H}) > wh(\tilde{D}, \tilde{H})$ reflecting the fact that overall education spending is higher under Θ_t, and ii) since $D = 0$ these agents have no education loans to repay. It follows from $I_{t+1}^c < 0$ that the package generates a positive surplus to future generations one that may be exploited to bring down pensions.[31] Finally, note that even if $G_t = \tilde{D}$ $\forall t$, private agents would have no incentive to borrow for education; moreover, I_{t+1}^c would still be negative.

Middle-aged in $t+k$, $k \geq 2$: From here on, the human capital externality term starts to evolve. The no-less utility condition for any generation $t+k$ is given

[31] Our use of the Pareto criterion follows a rich tradition in pension economics dating back, at least to Kotlikoff (1998). In that paper, Kotlikoff finds that privatization of the U.S. social security system (getting rid of its PAYG features) and financing the transition with a consumption tax can have a positive effect on welfare. The welfare of future generations rises, but older workers at the start of the transition experience significant welfare losses. He goes on to ask, "can a privatized system be Pareto improving?" Can the system "ensure that all agents who are alive at the start of the transition experience the identical level of utility during the transition that they would have experienced in the absence of privatization"?

by

$$wh\left(G_{t+k-1},H_{t+k-1}\right)+I_{t+k}\geq wh(\tilde{D},\tilde{H})-R\tilde{D}\Leftrightarrow I_{t+k}\geq I_{t+k}^{c}\equiv-A_{t+k}-R\tilde{D} \quad (18)$$

where

$$A_{t+k}\equiv wh\left(G_{t+k-1},H_{t+k-1}\right)-wh(\tilde{D},\tilde{H})\geq 0 \text{ and } H_{t+k}=h\left(G_{t+k-1},H_{t+k-1}\right).$$

Notice, in the absence of the HCE, $A_{t+k}=A_{t+k-1}$ and hence, both $I_t=0$ and $I_{t+k}=0$ $\forall k$ – this is true, for example, in the package implemented by BM. Here,

$$A_{t+k}\geq A_{t+k-1} \text{ for } G_{t+k-1}=G_{t+k}=G$$

due to the expanding human capital externality. The intuition why I_{t+k}^{c} continues to be negative (and why it may be possible to reduce the pension compensation for future generations) is that education produces an increase in human capital and thus income; hence, a given level of education offered for all generations will produce larger and larger income gains (relative to the CMA level) over time and cohorts.[32]

For future use, let us compute the lifetime present value to agents (at different dates) from their interaction with the government via Θ_t. For an agent born at $t-1$ (one who is middle-aged at t), this present value is given by

$$PV_{t-1}\equiv G_{t-1}-\frac{1}{R}\left(G_t+P_t\right)+\frac{1}{R^2}P_{t+1}=G_{t-1}+\frac{I_t}{R}. \quad (19)$$

Similarly,

$$PV_t=G_t-\frac{1}{R}\left(G_{t+1}+P_{t+1}\right)+\frac{1}{R^2}P_{t+2}=G_t+\frac{1}{R}I_{t+1}, \quad (20)$$

[32] Our analysis shares some superficial similarities with Pecchenino and Pollard (1997) who study a Diamond model, with uncertain (exogenous) survival to old age, a standard capital externality and an existing a PAYG pension system. In their setup, the scope for placing savings in life annuities is a policy parameter but full annuitization of savings is not optimal. They consider a policy reform – increasing the annuitization rate and decreasing the pension contribution rate (and thus the pension) – the initial situation being arbitrarily chosen with some annuitization rate and tax rate funding the public pension. This reform is evaluated under the Pareto criterion. They find that a gradual increase in the annuitization rate and a decrease in the pension contribution rate (tax) may lead to a Pareto improvement. This is possible since policy changes leading to more savings produces future gains via the capital externality. For an updated treatment of the Pechenino and Pollard (1997) idea, see Heijdra, Mierau, and Reijnders (2010).

and

$$PV_{t+k} = G_{t+k} - \frac{1}{R}\left(G_{t+k+1} + P_{t+k+1}\right) + \frac{1}{R^2}P_{t+k+2} = G_{t+k} + \frac{1}{R}I_{t+k+1} \quad (21)$$

where $I_{t+k+1} \equiv -G_{t+k+1} - P_{t+k+1} + \frac{1}{R}P_{t+k+2}$.

5.1 Pareto-improving pensions

In general, there may be an infinity of possible Pareto-improving paths characterized by $\{P_{t+j}\}_{j=0}^{\infty} \in \Theta_t$ that satisfy (16)-(18) – for an arbitrary sequence $\{G_{t+j}\}_{j=0}^{\infty}$ with $G_{t+j} \geq \tilde{D}$ for all j. Below, we analyze a specific package, $\Theta_t \equiv \{(G, P_{t+j})\}_{j=0}^{\infty}$ introduced at time t where, as before, $G \geq \tilde{D}$. The aim here is to characterize the Pareto-improving sequence $\{P_{t+j}\}_{j=0}^{\infty}$ associated with this package. Of special interest will be the non-monotonic nature of the sequence, P_{t+j}.

A small caveat needs to be recorded here. We have been using the term "Pareto-improving" to describe pensions that deliver the *same* (not higher) utility as under the CMA; perhaps, Pareto-neutral is a better terminology. Below we study settings in which these pensions, once they decline, fall forever. If we hadn't restricted pensions to be non-negative, they would eventually become negative. In other words, by restricting pensions to be non-negative, we are, in effect, distributing those "forbidden" welfare gains to future generations; it is in this sense, the path of pensions we study can legitimately be called "Pareto-improving".

Suppose all generations are given the minimum compensation (pension) that is needed to make them no worse/no better off than in the CMA. Then, using (16)-(18), the path of Pareto-improving pensions is given by

$$\begin{aligned}
&I_t^c = 0 \Leftrightarrow -G + \frac{1}{R}P_{t+1} = 0 \Leftrightarrow P_{t+1} = RG \\
&I_{t+1}^c = -A_{t+1} - R\tilde{D} \text{ where } A_{t+1} \equiv wh(G, \tilde{H}) - wh(\tilde{D}, \tilde{H}) \\
&\Leftrightarrow P_{t+2} = RG + RP_{t+1} - RA_{t+1} - R^2\tilde{D} \\
&I_{t+2}^c = -A_{t+2} - R\tilde{D} \text{ where } A_{t+2} \equiv wh(G, H_{t+1}) - wh(\tilde{D}, \tilde{H}) \\
&\Leftrightarrow P_{t+3} = R(G + P_{t+2}) - RA_{t+2} - R^2\tilde{D} \ ...
\end{aligned}$$

Clearly, the initial middle-aged generation needs to be compensated for bankrolling G. The amount of the pension must equal RG, the full opportunity cost of financing G, since they will see any countervailing income gains under the policy.

Subsequent generations may accept lower pensions because the opportunity cost of financing education is counterbalanced by future growth in income, fueled, in part, by the growth of the human capital externality. To see this more clearly, notice the path of Pareto-improving pensions k periods ahead is

$$P_{t+k} = R(G + P_{t+k-1}) - RA_{t+k-1} - R^2\tilde{D},$$

and similarly, $P_{t+k-1} = R(G + P_{t+k-2}) - RA_{t+k-2} - R^2\tilde{D}$. If the HCE was absent, $A_{t+k-2} = A_{t+k-1}$ would hold; then the path of pensions $P_{t+k} - P_{t+k-1} = R\left(P_{t+k-1} - P_{t+k-2}\right)$ would necessarily explode since $R > 1$. For BM, $G = \tilde{D}$ and $A_{t+1} = 0$ and hence $I^c_{t+1} = -R\tilde{D} \Leftrightarrow P_{t+2} = R\tilde{D}$; this combined with $I^c_t = 0 \Leftrightarrow P_{t+1} = R\tilde{D}$ implies Pareto-improving pensions in their world *cannot decline*; they would have to be *fixed forever at* $R\tilde{D}$.[33]

Proposition 3 *a) If $P_{t+k} < P_{t+k-1}$ for some j, then $P_{t+s} < P_{t+s-1}$ $\forall s \geq k$, i.e., if the path of Pareto-improving pensions is ever declining at some date, it is declining forever after.*
b) $P_{t+2} < P_{t+1}$ is not admissible, i.e., Pareto-improving pensions cannot start to decline rightaway.

Part (a) of Proposition 3 implies that the search for a declining sequence of Pareto-improving pensions ends with finding the first turning point. Once the level of pension compensation starts to decline, they do so, forever, until they reach zero. Part (b) of the proposition argues that $P_{t+2} > P_{t+1} > P_t = 0$ has to hold, meaning Pareto-improving pensions have to increase in generosity early on.

Some broad intuition for Proposition 3 is in order. When the human capital externality is present,

$$P_{t+k} - P_{t+k-1} = R\left(P_{t+k-1} - P_{t+k-2}\right) - R\left(A_{t+k-1} - A_{t+k-2}\right)$$

if $P_{t+k-1} - P_{t+k-2} < 0$, i.e., if the pension has started to decline, then $P_{t+k} - P_{t+k-1} < 0$ holds (pensions decline forever after) since $A_{t+k-1} - A_{t+k-2} > 0$ (because of the

[33] What BM show is that, in steady state, the CMA with private education, $\tilde{D}$, and return R is *equivalent* to an equilibrium with no private education and public education, $G = \tilde{D}$, together with a pension $P = R\tilde{D}$, financed by a tax levied on the middle-aged. They show this also holds outside of steady state if the level of public education is forever set equal to the level chosen by individuals with unfettered access to perfect education-loans markets at rate R.

contribution from the human capital externality).

The first round of pensions compensated the initial middle-aged for financing G since they received no such G when young. But in the following period, the then middle-aged have to finance both G and the pension RG to the initial middle-aged; hence they will need *more* pension compensation than what the initial middle-aged needed. This explains why pensions cannot start to decline rightaway.

Notice, as yet, the human capital externality hasn't kicked in – the last-mentioned middle-aged have parents who did not receive G when young. From the following period on, the human capital externality becomes active; these are the first generation whose parents received G when young and they are the first generation that sees an extra increase in their labor income due to this externality. Because of the assumed concavity in the human capital externality function, the marginal increase in labor income is the biggest it can ever be, for this generation; every subsequent generation will enjoy smaller and smaller marginal increases, i.e., $A_{t+j+1}-A_{t+j} \leq A_{t+j}-A_{t+j-1}$ since $A_{t+j+1} - A_{t+j} = w\left[h(G, H_{t+j}) - h(G, H_{t+j-1})\right] \geq 0$ and $H_{t+j} \geq H_{t+j-1}$ and $h_H \in (0, 1)$.

5.2 Declining pensions

Does this mean pensions, if they don't start declining by $t+3$, can never ever decline? Intuitively, it is clear that for the path of Pareto-improving pensions to exhibit a downward trend, the surplus – the gain in strength of the human capital externality – has to be sufficient to cover both the extra educational investment (since $G > \tilde{D}$) and the additional pension-finance burden. Below, we seek such a sufficient condition that generates the turning point alluded to in part (a) of Proposition 3. To that end, suppose the turning point happens at some date, $t+k$. In the appendix, we prove that

$$P_{t+k+1} - P_{t+k} = R^{k+1} \left[\underbrace{\left(G - \tilde{D}\right) - \frac{A_{t+1}}{R}}_{>0} - \sum_{j=1}^{k-1} R^{-j-1} \left(A_{t+j+1} - A_{t+j}\right) \right]. \quad (22)$$

Notice, the first term on the r.h.s. of (22) is positive, i.e., $R\left(G - \tilde{D}\right) > A_{t+1} \Leftrightarrow P_{t+2} > P_{t+1}$ which is true from part (b) of Proposition 3. Also note that $A_{t+j+1} - A_{t+j} \geq 0$. It follows from (22) that the path of pensions declines k periods down the road if the present value of increments to human capital, $\sum_{j=1}^{k-1} R^{-j-1}\left(A_{t+j+1} - A_{t+j}\right)$ is sufficiently large. Two factors are at play here – discounting and the growth in labor income. The externality releases big income increments in the near future and these have a large effect on the present value; increases in the distant future have a muted effect due to discounting. Hence it is possible that the turning point is as early as $t+3$, but it may come about later. In fact the present value is increasing in k, therefore the turning point does not *have* to be at $t+3$.[34] By way of contrast, notice that for Boldrin and Montes (2005), $G = \tilde{D}$, and hence $A_{t+j} = 0$ for all j, implying that $P_{t+k+1} - P_{t+k} = 0$ for all k, and as noted before, the pension at any date is fixed at $P = RG = R\tilde{D}$.

Below, we work out a sufficient condition for the path of pensions to decline, assuming the turning point comes at the earliest possible date – the third period ($k = 2$). That is, assume $P_{t+3} < P_{t+2}$ holds. It is easy to check that

$$P_{t+3} < P_{t+2} \Leftrightarrow A_{t+2} - A_{t+1} > R\left[R\left(G - \tilde{D}\right) - A_{t+1}\right] \tag{23}$$

The interpretation is clear. If the gain in human capital between period $t+1$ and $t+2$ is large enough to cover the burden of financing the past increase in pensions ($P_{t+2} - P_{t+1} = R\left[R\left[G - \tilde{D}\right] - A_{t+1}\right] > 0$), it is possible to decrease the pension to this generation and still keep them at least as well off as in the CMA. Define $\overline{G}_3$ to be the value of G which can be implemented such that (23) holds with equality, i.e.,

$$A_{t+2}+(R-1)A_{t+1} \equiv R^2\left(\overline{G} - \tilde{D}\right) \Leftrightarrow G < \overline{G}_3 \equiv \tilde{D}+\frac{A_{t+2}+(R-1)A_{t+1}}{R^2} \geq \tilde{D} \tag{24}$$

where the last inequality follows from $A_{t+2} = A_{t+1} = 0$ for $G = \tilde{D}$ and $A_{t+2} >$

[34] Suppose

$$\sum_{j=1}^{k-1} R^{-j-1}\left(A_{t+j+1} - A_{t+j}\right) < R^{-2}\left(A^{ss} - A_{t+1}\right)$$

where $A^{ss} \equiv \lim_{j\to\infty} A_{t+j}$. If

$$\left(G - \tilde{D}\right) - \frac{A_{t+1}}{R} > R^{-2}\left(A^{ss} - A_{t+1}\right)$$

then there does not exist a turning point.

$A_{t+1} > 0$ for $G > \tilde{D}$.[35] This means, for any $G < \overline{G}_3$, we are assured that the path of Pareto-improving pensions starts declining from $t+3$ on (and by part (a) of Proposition 3, declines forever after).

Proposition 4 *There exists a $\overline{G}_3 > \tilde{D}$ where $\overline{G}_3$ is defined in (24). Education policies $G \in \left[\tilde{D}, \overline{G}_3\right]$ are implementable and consistent with a path of pensions first increasing and then declining three periods after the policy is inaugurated.*

The implication is that there exists a package $\left\{\left(\overline{G}_3, P_{t+j}\right)\right\}_{j=0}^{\infty}$ consistent with overall education spending higher than under the CMA *as well as* a path of Pareto-improving pensions. Indeed, the path of Pareto-improving pensions is non-monotonic, rising early on in the transition and falling subsequently.[36]

5.3 Present value of the intergenerational compact

We proceed to compute the lifetime present value to agents (at different dates) from their interaction with the government via $\Theta_t \equiv \{(G, P_{t+j})\}_{j=0}^{\infty}$ where P_{t+j} satisfy the Pareto criterion. Restricting attention to Pareto-improving packages means I_t is replaced by I_t^c, I_{t+1} by I_{t+1}^c, and so on. Recall $I_t^c = 0$ and $I_{t+1}^c = -A_{t+1} - R\tilde{D}$; also, $I_{t+k+1}^c = -A_{t+k+1} - R\tilde{D}$. Noting $G_{t-1} = 0$, the lifetime present value calculations

[35] Using (22), one can define $\overline{G}_k$ by

$$\left(\overline{G}_k - \tilde{D}\right) - \frac{A_{t+1}\left(\overline{G}_k\right)}{R} - \sum_{j=1}^{k-1} R^{-j-1}\left(A_{t+j+1}\left(\overline{G}_k\right) - A_{t+j}\left(\overline{G}_k\right)\right) \equiv 0,$$

so that, the analog of (24) for a turning point at some $t+k$ is given by

$$\begin{aligned} \overline{G}_k &= \tilde{D} + \mathcal{F}\left(\overline{G}_k, k\right) \\ \mathcal{F}\left(\overline{G}_k, k\right) &\equiv \frac{1}{R^k}\left[A_{t+k}\left(\overline{G}_k\right) + (R-1)\sum_{j=1}^{k-1} R^{k-1-j} A_{t+j}\left(\overline{G}_k\right)\right] \end{aligned}$$

where $A_{t+j}(\overline{G}_k) \equiv wh(\overline{G}_k, h(\overline{G}_k, H_{t+j-1})) - wh(\tilde{D}, H\left(\tilde{D}\right))$. Since

$$\begin{aligned} F\left(\overline{G}_k, k\right) &> 0, F_G\left(\overline{G}_k, k\right) > 0, \\ F_{GG}\left(\overline{G}_k, k\right) &< 0 \text{ under generalized concavity assumption} \\ \text{and } F_k\left(\overline{G}_k, k\right) & \quad 0 \end{aligned}$$

$\overline{G}_k$ is well defined for all k, and $\overline{G}_k > \tilde{D}$.

[36] It bears emphasis here that the phasing out of PAYG pensions is possible assuming period by period budget balance; that is, the argument does not rely on the public sector being able to use capital markets in a more sophisticated way than private agents can.

presented in (19)-(21) are adjusted as described below.

$$PV_{t-1} = 0.$$

Then,

$$PV_t = G + \frac{1}{R} I^c_{t+1} = \frac{RG - \left(A_{t+1} + R\tilde{D}\right)}{R}.$$

We know from Proposition 3 that $P_{t+2} > P_{t+1}$ has to hold. This implies $RG - \left(A_{t+1} + R\tilde{D}\right) > 0 \Leftrightarrow PV_t > 0$. Note that

$$PV_{t+k+1} - PV_{t+k} = \frac{1}{R}\left(I^c_{t+k+2} - I^c_{t+k+1}\right) = \frac{1}{R}\left(-A_{t+k+2} + A_{t+k+1}\right) < 0 \Longleftrightarrow PV_{t+k} > PV_{t+k+1}$$

or a declining sequence of present values. Note, in the long run, when the pension has been eliminated,

$$PV_\infty = G - \frac{1}{R} G = G\left(\frac{R-1}{R}\right) > 0$$

The upshot is that from the point of view of the agents, and under the Pareto criterion, the present value of their interaction with the public sector is highest for the first generation of young that receives public education; subsequent generations see a decline in the net expansion of their budget sets, but the present value itself remains *positive* forever. These present value calculations serve to highlight the conformity of our results with the information contained in Figures 1 and 2. They underscore the fact that the "borrowing" component of the intergenerational arrangement is dominating, suggesting welfare gains are associated with the compact.

To sum up: there may be downstream welfare gains from the public provision of education even when capital markets are complete. The trouble is, these gains may never materialize if some initial generations have to suffer to bring them online. We show it is possible to implement upto a certain level of publicly-funded education that generates higher levels of human capital than under complete education-loan markets, and yet, leaves every current and future generation unhurt or even better off. We characterize the optimal path of pensions that must accompany such a education-spending policy. We show that if the human capital externality is strong enough, the gains early on are enough to support an eventual phasing-out of public pensions with possible replacement by fully-funded pension schemes.

6 Conclusion

The intergenerational welfare state has two arms, providing for the young (education) and the old (pensions), all of which is financed by taxes levied on the middle-aged. In this paper, we show it is important to consider both arms jointly, and that there is a fundamental difference between the forward and backward looking part of the intergenerational compact. In a dynamically efficient economy, the forward part (the borrowing part) is associated with welfare gains, while the backward part (the saving part), needed to compensate tax payers for financing education for the young, is associated with welfare losses. However, both are needed if welfare gains are to be released without any generations suffering along the way. We have shown via the intergenerational compact it is possible to implement higher levels of education than what is achievable via complete markets, and that this is associated with welfare gains. Importantly, the transition path is consistent with the pension arm withering away over time. The welfare gains associated with the intergenerational compact are intimately associated with its borrowing aspects, and in accordance with empirical evidence, we find the expected net present value to the individual from the package is always non-negative, and strictly positive when the transition has ended. The higher level of education leads to higher output and thus higher incomes, implying the welfare state is associated with an improvement in material well-being. We show this is possible to bring about without harming any generations.

These findings have several important implications. It suggests the traditional interpretation of PAYG pensions as merely an unnecessary and costly gift from future generations to the inaugural generation is potentially misleading. Pensions have an important compensatory role to play in the phasing in of welfare-improving public sector activities (education). The compensatory role of pensions can dwindle over time, and this allows them to be phased out. The investment role of the welfare state remains important throughout. For us, pension reform is not a retrenchment of the welfare state; instead, it suggests pensions may have served their purpose.

Our analysis has been cognizant of the historical evolution of the welfare state. After all, historical experience from nineteenth-century Europe, and even the United States – see Tanzi and Schuknecht (2000) – suggests that establishment of institutions of public education preceded, sometimes by more than two generations, those of public pensions. Our analysis may also be interesting from the perspective of governments in fast-developing countries, such as China or Korea, which are currently contemplating introducing PAYG pensions even as they continue to invest heavily in public education.

Our analysis has stayed away from studying alternative schemes that boost or help private education spending, such as a direct subsidy on education spending or a interest subsidy on the costs of education-related borrowing. Any such scheme would require public funding paid for by taxes on the middle-aged. As such, any attempt to introduce these under the Pareto criterion would presumably face similar implementation hurdles as raised here. In a similar vein, financing via distortionary taxes could also be considered. We leave these extensions to future research.

Appendix

A The borrowing and saving parts of the social compact

We present some details regarding the intergenerational compact. The key input is the age dependent net-transfer (contribution) between the individual and the public sector, n_a, where a denotes age. Here n_a is the net transfer to the agent and $-n_a$ is the net contribution by the agent to the public sector. By net transfer is understood the value of services provided in kind and various cash transfers net of taxes paid at various ages. The numbers presented below are averages for a representative person of a given age. Note intergenerational (re)distribution is not considered.

First some definitions. Let π_a denote the population share of given age a. The the intergenerational compact is financially viable provided

$$-\sum_{a=0}^{A}\pi_a n_a \geq 0 \Leftrightarrow \sum_{a=0}^{A}\pi_a n_a \leq 0$$

where A is the maximal possible age. From an individual's perspective, the expected present value PV for a newborn accepting the contract is

$$PV_0^{IGC} = \sum_{a=0}^{A}\left(\frac{1}{1+r}\right)^a \pi_a n_a$$

where r is the discount rate. Note that the expected present value of the contributions made by a newborn to the public sector PVP is

$$PVP_0^{IGC} = -PV_0^{IGC} = -\sum_{j=a}^{A}\left(\frac{1}{1+r}\right)^a \pi_a n_a$$

assuming the same discount rate for individuals and for the public sector.

We wish to make the point that the financial viability or budget constraint $-\sum_{a=0}^{A}\pi_a n_a \geq 0 \not\Leftrightarrow PVP_0^{IGC} \geq 0$ ($PV_0^{IGC} \leq 0$).This is most easily demonstrated in a standard two-period overlapping generations structure, with young and old overlapping. Let generation born at t be of size N_t where $N_t = (1+n)N_{t-1}$. The survival rate is, for simplicity, assumed to be 100%. Consider a scheme where all from generation $t-1$ (the young) contribute an amount T_t which finances a transfer P_t to generation t (the old). The transfer is determined as

$$P_t N_{t-1} = T_t N_t \Rightarrow P_t = T_t(1+n)$$

Note this is a PAYG scheme, and hence it it is financially viable. Assume that this is so for all generations, i.e., we consider a stationary scheme ($P_t = P$ and $T_t = T$ for all t). The present value of the net contributions to this scheme seen from the government perspective is

$$PVP_0^{IGC} = T - \frac{P}{1+r} = T - \frac{T(1+n)}{(1+r)} = T\left(\frac{r-n}{1+r}\right) > 0 \text{ for } r > n$$

From an individual perspective we have the reverse

$$PV_0^{IGC} = -T + \frac{T(1+n)}{(1+r)} = T\left(\frac{n-r}{1+r}\right) < 0 \text{ for } r > n$$

where $r > n$ ensures dynamic efficiency. It follows, in this scheme – which by construction is financially viable – each individual in present value terms contributes more than she receives from the social contract. Therefore, from an individual perspective this scheme has a *negative* present value. This is another way of saying it has a return (n) below the market return (r).

Consider now the opposite scheme where all from generation $t-1$ (the old) contribute an amount T_t going to finance a transfer P_t to generation t (the young). In this case, the transfer is determined as $T_t = P_t(1+n)$ and for a stationary scheme, the present value of the net contributions from the public sector perspective is

$$PVP_0^{IGC} = -P + \frac{T}{1+r} = -P + \frac{P(1+n)}{1+r} = -P\left(\frac{r-n}{1+r}\right) < 0 \text{ for } r > n$$

and the reverse from an individual perspective:

$$PV_0^{IGC} = P - \frac{T}{1+r} = P\left(\frac{r-n}{1+r}\right) > 0 \text{ for } r > n$$

Such a scheme – which by construction is financially viable – allows individuals in expected present value terms to contribute less than is received from the social contract. From an individual perspective, this interaction with the public sector yields a positive present value. Moreover, it is a borrowing scheme, something is received first, and later paid back. Since the interest (n) is lower than the market interest rate (r), it is favorable to borrow in this scheme. This suggests that an EP package which is frontloaded, i.e., where the borrowing part is dominant, is potentially welfare improving.

B Assumptions on the human capital technology

Since education-loan markets are complete and private agents adjust their borrowings one-for-one against government top-up spending, an arbitrary $\{G_t\}$ package would, at times correspond to a $\{D_t\}$ given by $wh_E(D_{t-1} + G_{t-1}, H_{t-1}) = R$, and when G is big enough, to $D_t = 0$ with $wh_E(G_{t-1}, H_{t-1}) < R$. It is readily apparent that this "two-pronged" definition of D would pose serious analytical challenges. To keep things more manageable, choose G to be high enough so that $E = G$ and $D = 0$ at the inaugural date; private education is fully crowded out on impact. This, however, is not enough. Since the HCE increases the marginal product of human capital investments, further along, a private incentive to undertake education expenses may get reinstated. We further need assurance that private education will stay fully crowded out along the entire transition. More formally, we require $wh_E(E, H_{t-1}) < R$ for $G > \tilde{D}$ but additionally, we need $wh_E(E, H_{t+k}) < R$ for all $k \geq 0$. Since $H_{t+k} \geq H_{t+k-1}$ for all k, and $h_{EH} > 0$ is assumed, the latter condition may not hold.

Our initial situation has

$$wh_E(\tilde{D}, H(\tilde{D})) = R$$

Consider any level of education $E > \tilde{D}$. Then $h_{EE} < 0$ implies

$$wh_E(E, H(\tilde{D})) < R.$$

Monotonicity implies $H_{t+k} \leq H(E)$ for all $k \geq 0$. Since $h_{EH}(\cdot) > 0$, it follows that

$$h_E(E, H(E)) \geq h_E(E, H_{t+k}) \text{ for all } k \geq 0.$$

Hence, if

$$h_{EE}(E, H(E)) + h_{EH}(E, H(E))H_E(E) < 0$$

holds – Assumption 3.1 – it follows that

$$wh_E(E, H(E)) < R$$

and hence

$$wh_E(E, H_{t+k}) < R \quad \text{for all } k \geq 0.$$

As an example, consider the special functional form

$$H_t = h\left(E_{t-1}, H_{t-1}\right) = E_{t-1}^{\alpha} H_{t-1}^{\beta}.$$

Then,

$$h_E\left(\cdot\right) = \alpha E_{t-1}^{\alpha-1} H_{t-1}^{\beta};\; h_{EE}\left(\cdot\right) = \alpha\left(\alpha - 1\right) E_{t-1}^{\alpha-2} H_{t-1}^{\beta};\; h_{EH}\left(\cdot\right) = \alpha\beta E_{t-1}^{\alpha-1} H_{t-1}^{\beta-1}.$$

Steady state human capital satisfies

$$H = E^{\frac{\alpha}{1-\beta}} \Longrightarrow H_E = \frac{\alpha}{1-\beta} E^{\frac{\alpha}{1-\beta}-1}.$$

Then, routine algebra verifies that

$$h_{EE}(E, H(E)) + h_{EH}(E, H(E))H_E(E) = \left[\alpha(\alpha - 1) + \alpha\beta\frac{\alpha}{1-\beta}\right] E^{\frac{\alpha}{1-\beta}-2}.$$

It follows that

$$\alpha(\alpha - 1) + \alpha\beta\frac{\alpha}{1-\beta} = \alpha\frac{\alpha+\beta-1}{1-\beta} < 0$$

holds for $\alpha + \beta < 1$.

C Implementing the Golden rule

We have shown it is possible to implement education levels that generate higher levels of human capital than under complete education-loan markets. Specifically, there is a range of public-education spending levels, from a lower bound of that achievable under complete private markets to the upper bound level of the Golden rule; our results thus far indicate that any level in the *interior* of this range may be implementable. But, can the Golden rule itself be implemented? That is, starting from a CMA steady state, can the government implement (under the Pareto criterion) a education-spending jump from $\tilde{D}$ to G^{gr} at date t and thereafter hold it at that level? Below, we show this is not possible if pensions must fall at its earliest possible date.

Proposition 5 *The Golden rule level of education spending, defined in (15), cannot be implemented if pensions must start to decline from the third period on.*

Proof. The Golden rule satisfies $wh_E(G^{GR}, H(G^{GR})) \equiv 1$. Is this level of G among the packages that are implementable? That is, is $G^{GR} \in \left[\tilde{D}, \overline{G}_3\right]$? For this to hold, we need

$$G^{GR} < \tilde{D} + F(G^{GR}) < \tilde{D} + \frac{A^{ss}(G^{GR})}{R}$$

where

$$A^{ss}(G^{GR}) \equiv wh(G^{GR}, H(G^{GR})) - wh_E(\tilde{D}, H(\tilde{D}))$$

Using a second-order Taylor approximation, we have

$$\begin{aligned} wh(G^{GR}, H(G^{GR})) - wh_E(\tilde{D}, H(\tilde{D})) &= wh_E(\tilde{D}, H(\tilde{D}))\left[G^{GR} - \tilde{D}\right] + \frac{1}{2}h_{EE}(\tilde{D}, H(\tilde{D}))\left[G^{GR} - \tilde{D}\right]^2 \\ &= R\left[G^{GR} - \tilde{D}\right] + \frac{1}{2}h_{EE}(\tilde{D}, H(\tilde{D}))\left[G^{GR} - \tilde{D}\right]^2 \end{aligned}$$

Then,

$$\begin{aligned}G^{GR} &< \tilde{D}+\frac{A^{ss}(G)}{R} \Leftrightarrow G^{GR} < \tilde{D}+\left[G^{GR}-\tilde{D}\right]+\frac{1}{2R}h_{EE}(\tilde{D},H(\tilde{D}))\left[G^{GR}-\tilde{D}\right]^2 \\ \Leftrightarrow\ & 0<\frac{1}{2R}h_{EE}(\tilde{D},H(\tilde{D}))\left[G^{GR}-\tilde{D}\right]^2<0\end{aligned}$$

a contradiction. The Golden rule level of education cannot be implemented by this package. ■

Of course, this doesn't preclude G^{gr} from being implementable if pensions start to decline at *some* date $t+k$ where $k>2$. These sorts of issues are investigated using numerical methods below. To that end, consider a special functional form for human capital accumulation:

$$H_t = h\left(E_{t-1},H_{t-1}\right)=E_{t-1}^{\alpha}H_{t-1}^{\beta};\quad \alpha,\beta\in(0,1)\,;\ \alpha+\beta<1$$

and $E_{t-1}\equiv D_{t-1}+G$.[37] For this specification, it is easy to verify that

$$\tilde{D}=\left[\frac{R}{w\alpha}\right]^{\frac{1-\beta}{\alpha+\beta-1}} \text{ and } G^{gr}=\left[\frac{1-\beta}{\alpha}\frac{1}{w}\right]^{\frac{1-\beta}{\alpha+\beta-1}},$$

and hence

$$\frac{G^{gr}}{\tilde{D}}=\left[\frac{1-\beta}{R}\right]^{\frac{1-\beta}{\alpha+\beta-1}}>1 \text{ since } 1-\beta<R \text{ and } \frac{1-\beta}{\alpha+\beta-1}<0.$$

This clearly establishes that the difference between the education level in the CMA and at the Golden Rule is driven by the intergenerational human capital externality (β) and the gross return (R) cf, the aforediscussed borrowing argument. Figure 3 illustrates the path of Pareto-improving pensions (for a high and a low value of β) consistent with implementation of the Golden rule level of education. In one case (High Beta) the pension is quickly phased out, while in the other (Low Beta) the pension stays fairly constant for a while until it is eventually phased out.

[37] This formulation obviously violates $h\left(0,0\right)>0$. It is trivial to add in a small constant to rectify this problem.

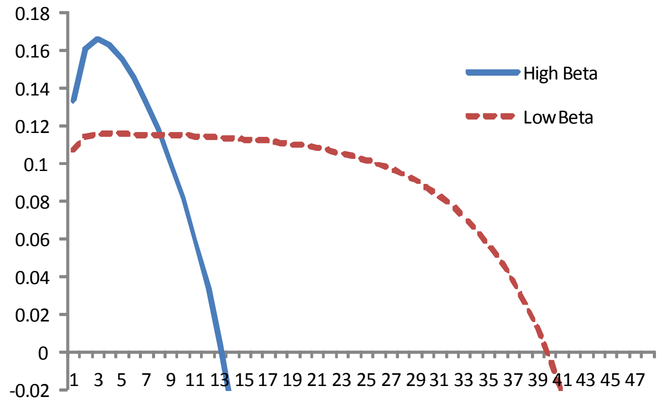

Figure 3: Pension paths implementing the Golden Rule level of education

Note: The vertical axis measures P under the Pareto criterion and the horizontal axis is t. Computed for the following parameter values: $\alpha = 0.1$, $\beta = 0.4$ (High Beta), $\beta = 0.1$ (Low Beta), $w = 1$, $R = 1.15$.

A change in β has three effects: it changes the education level in the CMA ($\tilde{D}$), the Golden rule level of education (G^{gr}), and the speed of increments in human capital (and hence, the pension path under the Pareto criterion, cf. (22). In general, it is ambiguous whether an increase in β makes it more or less easy to implement the Golden-Rule level of education.

D An extension: private and public education as complements

All along, we have worked under the assumption that private and public spending on education are perfect substitutes. Aside from the analytical convenience, such an assumption allows public spending to crowd out and drive private spending to the zero corner. In this extension, we allow for the possibility that private and public spending on education are complements implying that private education is not fully crowded out. The goal is to (numerically) study the possibility the Golden rule is implementable even in this case.

To that end, consider a human capital equation of the following form:

$$H_t = h(D_t, G_t, H_t) = (D_{t-1} + \underline{D})^{\gamma} (G_{t-1} + \underline{G})^{\theta} H_{t-1}^{\beta}$$

where $\gamma + \theta + \beta < 1, \gamma > 0, \theta > 0, \beta > 0$, $\underline{D} > 0$ and $\underline{G} > 0$. This specification implies private (D) and public (G) educational inputs are complements, that is, the higher either input, the higher the marginal product of the other. The optimal private level of education is given as

$$w\gamma (D_{t-1} + \underline{D})^{\gamma-1} (G_{t-1} + \underline{G})^{\theta} H_{t-1}^{\beta} = R.$$

In the CMA (marked by $\tilde{}$) where there is no public educational input ($G = 0$), the optimal level of private education is

$$\tilde{D} + \underline{D} = \left(\frac{R}{w\gamma}\right)^{\frac{1}{\gamma-1}} (\underline{G})^{\frac{-\theta}{\gamma-1}} \left(\tilde{H}\right)^{\frac{-\beta}{\gamma-1}}.$$

It is seen that a higher level of public education induces a higher level of private education ($\frac{-\theta}{\gamma-1} > 0$), i.e., the two forms of education are complements.

Since the steady-state human capital stock can be written as $\tilde{H} = \left(\tilde{D} + \underline{D}\right)^{\frac{\gamma}{1-\beta}} (\underline{G})^{\frac{\theta}{1-\beta}}$, we have

$$\left(\tilde{D} + \underline{D}\right) = \left(\frac{R}{w\gamma}\right)^{\frac{1-\beta}{\gamma+\beta-1}} (\underline{G})^{\frac{-\theta}{\gamma+\beta-1}}$$

To determine the Golden Rule level of education, we first find the steady state level of human capital for a given level of public education. Since $H = (D + \underline{D})^{\frac{\gamma}{1-\beta}} (G + \underline{G})^{\frac{\theta}{1-\beta}}$, using the first order condition for private education, it follows that

$$H = \left(\frac{R}{w\gamma}\right)^{\frac{\gamma}{\gamma+\beta-1}} (G + \underline{G})^{\frac{\theta}{1-\gamma-\beta}}.$$

The Golden Rule is characterized by $w\frac{\partial H}{\partial G} = 1$, and hence

$$(G^{GR} + \underline{G}) = \left(\frac{1-\gamma-\beta}{w\theta}\right)^{\frac{1-\gamma-\beta}{\theta+\gamma+\beta-1}} \left(\frac{R}{w\gamma}\right)^{\frac{\gamma}{\theta+\gamma+\beta-1}}.$$

When introducing public education, the economy will follow a trajectory where the human capital stock evolves according to

$$H_t = \left(\frac{R}{w\gamma}\right)^{\frac{\gamma}{\gamma-1}} (G_{t-1} + \underline{G})^{\frac{\theta}{1-\gamma}} H_{t-1}^{\frac{\beta}{1-\gamma}},$$

where the first order condition for private education has been used to determine the private educational input.

Finally, note that the pension satisfying the Pareto criterion evolves differently when educational inputs are complements, since public consumption no longer fully crowds out private education. This in turn also implies that incomes are affected by changes in private education, and hence, debt repayment. Analogous to the procedure in the main text, considering the middle-aged in some period $t + k$, the no-less utility condition for any generation $t + k$ is given by

$$wH_{t+k} - RD_{t+k-1} + I_{t+k} \geq w\tilde{H} - R\tilde{D} \Leftrightarrow I_{t+k} \geq I^c_{t+k} \equiv -A^c_{t+k} - R\tilde{D}$$

where

$$A^c_{t+k} \equiv wh\left(G_{t+k-1}, D_{t+k-1}, H_{t+k-1}\right) - wh\left(\tilde{D}, \tilde{H}\right) + RD_{t+k-1} \geq 0 \text{ and } H_{t+k} = h\left(G_{t+k-1}, D_{t+k-1}, H_{t+k}\right.$$

If G is implemented, the pension evolves according to

$$P_{t+k} = R(G + P_{t+k-1}) - RA^c_{t+k-1} - R^2\tilde{D},\ k \geq 2$$

while the pension in $t+1$ as before is $P_{t+1} = RG$.

The figure below shows a path for the pension implementing the Golden Rule level of education starting out from the CMA. Note that the introduction of public education due to the complementarity increases private education and thus private borrowing. As shown in the expression above, the changed borrowing affects the equal pay-off condition under the Pareto criterion and thus the path for the pension.

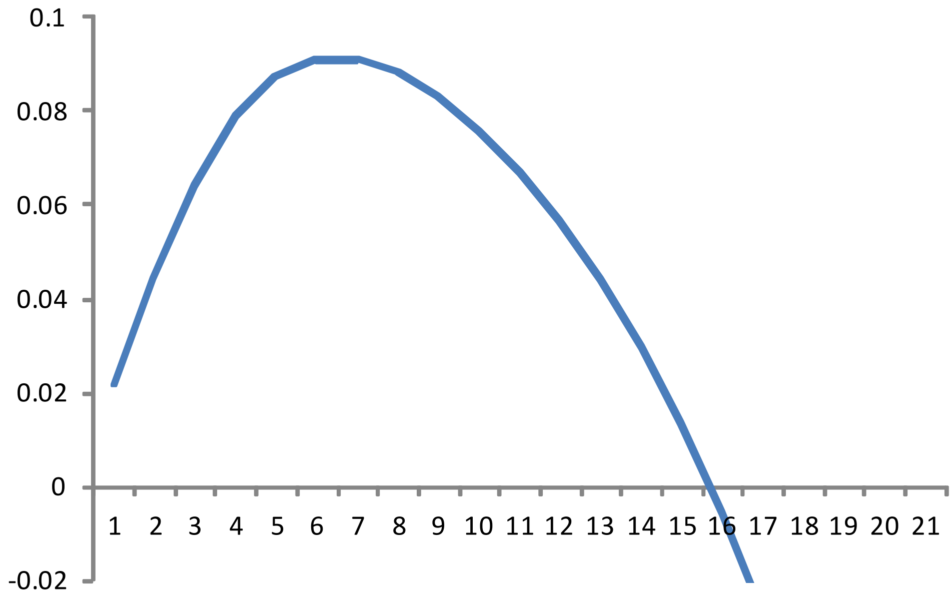

Figure 4: Pension paths implementing the Golden Rule level with complementarity between private and public education

Note: The vertical axis measures P under the Pareto criterion and the horizontal axis is t. Computed for the following parameter values: $\gamma = \delta = 0.2$, $\beta = 0.47$, $w = 1$, $R = 1.15$, $\underline{D} = \underline{G} = 0.0001$.

E Proof of Proposition 3

a) Using $P_{t+j} = R(G + P_{t+j-1}) - RA_{t+j-1} - R^2\tilde{D}$, it is easy to derive $P_{t+j} - P_{t+j-1} = R(P_{t+j-1} - P_{t+j-2}) - R(A_{t+j-1} - A_{t+j-2})$. Since $(A_{t+j} - A_{t+j-1}) \geq 0$ for all j, it

follows that if $(P_{t+j-1} - P_{t+j-2}) < 0$, then $P_{t+j} - P_{t+j-1} < 0$. Forward iteration implies, that if $P_{t+j} < P_{t+j-1}$ for some j, then $P_{t+s} < P_{t+s-1}$ $\forall s \geq j$.
b) Consider the possibility

$$P_{t+2} < P_{t+1} \Leftrightarrow R(G + RG) - RA_{t+1} - R^2\tilde{D} < RG \Leftrightarrow RG - A_{t+1} - R\tilde{D} < 0$$

which, in turn, requires $A_{t+1} > R\left(G - \tilde{D}\right)$.Recall $\tilde{H} \equiv H(\tilde{D})$ and that $wh_E(\tilde{D}, \tilde{H}) = R$. Making a second-order Taylor-approximation of $h(G, \tilde{H})$ around $(\tilde{D}, \tilde{H})$ we have

$$h(G, \tilde{H}) \simeq h(\tilde{D}, \tilde{H}) + h_E(\tilde{D}, \tilde{H})(G - \tilde{D}) + \frac{1}{2}h_{EE}(\tilde{D}, \tilde{H})(G - \tilde{D})^2.$$

Hence

$$A_{t+1} = wh(G, \tilde{H}) - wh(\tilde{D}, \tilde{H}) \simeq R(G - \tilde{D}) + \frac{w}{2}h_{EE}(\tilde{D}, \tilde{H})(G - \tilde{D})^2$$

where it is used that $wh_E(\tilde{D}, \tilde{H}) = R$. Then, $A_{t+1} > R\left(G - \tilde{D}\right)$ reads

$$R(G-\tilde{D}) + \frac{w}{2}h_{EE}(\tilde{D}, \tilde{H})(G-\tilde{D})^2 > R\left(G - \tilde{D}\right) \Longrightarrow \frac{w}{2}h_{EE}(\tilde{D}, \tilde{H})(G-\tilde{D})^2 > 0.$$

This last inequality is violated since $h_{EE}(\tilde{D}, \tilde{H}) < 0$.

F Derivation of Eq. (22)

We have that

$$\begin{aligned}
P_{t+1} &= RG \\
P_{t+2} &= R(G + RG) - RA_{t+k-1} - R^2\tilde{D} = RG + R^2\left(G - \tilde{D}\right) - RA_{t+1} \\
P_{t+3} &= RG + R\left[RG + R^2\left(G - \tilde{D}\right) - RA_{t+1}\right] - RA_{t+2} - R^2\tilde{D} \\
&= RG + R^2\left(G - \tilde{D}\right) + R^3\left(G - \tilde{D}\right) - R^2A_{t+1} - RA_{t+2} \\
P_{t+4} &= RG + R^2\left(G - \tilde{D}\right) + R^3\left(G - \tilde{D}\right) + R^4\left(G - \tilde{D}\right) - R^3A_{t+1} - R^2A_{t+2} - RA_{t+3}.... \\
P_{t+k} &= RG + R^2\left(G - \tilde{D}\right)\sum_{j=0}^{k-2} R^j - R\sum_{j=1}^{k-1} R^{k-1-j}A_{t+j}
\end{aligned}$$

Leading this equation one period and subtracting yields

$$
\begin{aligned}
&P_{t+k+1} - P_{t+k} = RG + R^2\left(G - \tilde{D}\right)\sum_{j=0}^{k-1} R^j - R\sum_{j=1}^{k} R^{k-j}A_{t+j} \\
&- \left(RG + R^2\left(G - \tilde{D}\right)\sum_{j=0}^{k-2} R^j - R\sum_{j=1}^{k-1} R^{k-1-j}A_{t+j}\right) \\
&= R^2\left(G - \tilde{D}\right)\left[\sum_{j=0}^{k-1} R^j - \sum_{j=0}^{k-2} R^j\right] + R\left[\sum_{j=1}^{k-1} R^{k-1-j}A_{t+j} - \sum_{j=1}^{k} R^{k-j}A_{t+j}\right].
\end{aligned}
$$

Noting that $\sum_{j=0}^{k-1} R^j - \sum_{j=0}^{k-2} R^j = R^{k-1}$,we have

$$
P_{t+k+1} - P_{t+k} = R^{k+1}\left(G - \tilde{D}\right) + R\left[\sum_{j=1}^{k-1} R^{k-1-j}A_{t+j} - \sum_{j=1}^{k} R^{k-j}A_{t+j}\right].
$$

Furthermore we have

$$
\begin{aligned}
&\sum_{j=1}^{k-1} R^{k-1-j}A_{t+j} - \sum_{j=1}^{k} R^{k-j}A_{t+j} \\
= \;& R^{k-2}A_{t+1} + R^{k-3}A_{t+2} + R^{k-4}A_{t+3} + \ldots + A_{t+k-1} - R^{k-1}A_{t+1} - R^{k-2}A_{t+2} - R^{k-3}A_{t+3} - \ldots - A \\
= \;& -R^{k-1}A_{t+1} + R^{k-2}A_{t+1} - R^{k-2}A_{t+2} + R^{k-3}A_{t+2} - R^{k-3}A_{t+3} + R^{k-4}A_{t+3} - \ldots \\
= \;& -R^{k-1}A_{t+1} + R^{k-2}\left(A_{t+1} - A_{t+2}\right) + R^{k-3}\left(A_{t+2} - A_{t+3}\right) + \ldots + \left(A_{t+k-1} - A_{t+k}\right) \\
= \;& -R^{k-1}A_{t+1} + \sum_{j=1}^{k-1} R^{k-j-1}\left(A_{t+j} - A_{t+j+1}\right).
\end{aligned}
$$

Hence, it follows that

$$
\begin{aligned}
&P_{t+k+1} - P_{t+k} = R^{k+1}\left(G - \tilde{D}\right) + R\left[\sum_{j=1}^{k-1} R^{k-1-j}A_{t+j} - \sum_{j=1}^{k} R^{k-j}A_{t+j}\right] \\
&= R^{k+1}\left(G - \tilde{D}\right) - R^kA_{t+1} + R\sum_{j=1}^{k-1} R^{k-j-1}\left(A_{t+j} - A_{t+j+1}\right) \\
&= R^{k+1}\left[\left(G - \tilde{D}\right) - \frac{A_{t+1}}{R} - \frac{1}{R^k}\sum_{j=1}^{k-1} R^{k-j-1}\left(A_{t+j+1} - A_{t+j}\right)\right].
\end{aligned}
$$

G Proof of Proposition 4

It follows from (24) that

$$\begin{aligned}\overline{G}_3 &= \tilde{D} + \mathcal{F}(\overline{G}_3) \text{ where} \\ \mathcal{F}(\overline{G}_3) &\equiv \frac{A_{t+2}(\overline{G}_3) + (R-1)A_{t+1}(\overline{G}_3)}{R^2}\end{aligned}$$

Notice

$$\begin{aligned}\mathcal{F}(\tilde{D}) &= 0 \\ \mathcal{F}_G(\overline{G}_3) &= \frac{1}{R^2}\left[\frac{\partial A_{t+2}(\overline{G}_3)}{\partial \overline{G}_3} + (R-1)\frac{\partial A_{t+1}(\overline{G}_3)}{\partial \overline{G}_3}\right] \\ \mathcal{F}_{GG}(\overline{G}_3) &= \frac{1}{R^2}\left[\frac{\partial A^2_{t+2}(\overline{G}_3)}{\partial \overline{G}^2_3} + (R-1)\frac{\partial A^2_{t+1}(\overline{G}_3)}{\partial \overline{G}^2_3}\right]\end{aligned}$$

Using $A_{t+1}(\overline{G}_3) = wh(\overline{G}_3, \tilde{H}) - wh(\tilde{D}, \tilde{H})$ and $A_{t+2}(\overline{G}_3) = wh(\overline{G}_3, h(\overline{G}_3, \tilde{H})) - wh(\tilde{D}, \tilde{H})$, we get

$$\begin{aligned}\frac{\partial A_{t+1}(\overline{G}_3)}{\partial \overline{G}_3} &= wh_E(\overline{G}_3, \tilde{H}) > 0 \\ \frac{\partial A_{t+2}(\overline{G}_3)}{\partial \overline{G}_3} &= wh_E(\overline{G}_3, h(\overline{G}_3, \tilde{H})) + wh_H(\overline{G}_3, h(\overline{G}_3, \tilde{H}))h_E(\overline{G}_3, \tilde{H}) > 0 \\ \frac{\partial A^2_{t+1}(\overline{G}_3)}{\partial \overline{G}^2_3} &= wh_{EE}(\overline{G}_3, \tilde{H}) < 0 \\ \frac{\partial A^2_{t+2}(\overline{G}_3)}{\partial \overline{G}^2_3} &= w\left[\begin{array}{c} h_{EE}(\overline{G}_3, h(\overline{G}_3, \tilde{H})) + h_{EH}(\overline{G}_3, h(\overline{G}_3, \tilde{H}))h_E(\overline{G}_3, \tilde{H}) \\ + \left[h_{HE}(\overline{G}_3, h(\overline{G}_3, \tilde{H})) + h_{HH}(\overline{G}_3, h(\overline{G}_3, \tilde{H}))h_E(\overline{G}_3, \tilde{H})\right] h_E(\overline{G}_3, \tilde{H}) \\ + h_H(\overline{G}_3, h(\overline{G}_3, \tilde{H}))h_{EE}(\overline{G}_3, \tilde{H}) \end{array}\right] < 0\end{aligned}$$

Notice that the last inequality follows from the fact that $\frac{\partial^2 H_{t+2}}{\partial E^2} < 0$, cf assumption made above.

It follows that $\mathcal{F}(.)$ is increasing and concave. If we can show that the slope of the $\mathcal{F}$-function is larger than 1 at $G = \tilde{D}$, then concavity of $\mathcal{F}(.)$ would ensure the existence of a $\overline{G}_3$. To that end, note

$$\begin{aligned}&\mathcal{F}_G(\tilde{D}) = \frac{1}{R^2}\left[\frac{\partial A_{t+2}(\tilde{D})}{\partial \overline{G}} + (R-1)\frac{\partial A_{t+1}(\tilde{D})}{\partial \overline{G}}\right] > 1 \\ &\Longrightarrow \frac{1}{R^2}\left[wh_E(\tilde{D}, h(\tilde{D}, \tilde{H})) + wh_H(\tilde{D}, h(\tilde{D}, \tilde{H}))h_E(\tilde{D}, \tilde{H}) + (R-1)wh_E(\tilde{D}, \tilde{H})\right] > 1 \\ &\Longrightarrow \frac{1}{R^2}\left[wh_H(\tilde{D}, h(\tilde{D}, \tilde{H}))h_E(\tilde{D}, \tilde{H}) + Rwh_E(\tilde{D}, \tilde{H})\right] > 1.\end{aligned}$$

The last inequality implies

$$
\begin{aligned}
& w h_H(\tilde{D}, h(\tilde{D}, \tilde{H})) h_E(\tilde{D}, \tilde{H}) + R w h_E(\tilde{D}, \tilde{H}) > R^2 \\
& \Longrightarrow w \left[h_H(\tilde{D}, h(\tilde{D}, \tilde{H})) + R \right] h_E(\tilde{D}, \tilde{H}) > R^2 \\
& \Longrightarrow h_H(\tilde{D}, h(\tilde{D}, \tilde{H})) + R > R \\
& \Longrightarrow h_H(\tilde{D}, h(\tilde{D}, \tilde{H})) > 0
\end{aligned}
$$

which holds (notice, we have used $w h_E(\tilde{D}, \tilde{H}) = R$). This proves there exists $\overline{G}_3 > \tilde{D}$ and it is bounded above.

References

[1] Aaron, Henry. 1966. "The social insurance paradox," *Canadian Journal of Economics and Political Science* 32(3): 371–374.

[2] Andolfatto, D., Gervais, M., 2006. Human capital investment and debt constraints. *Review of Economic Dynamics* 9(1), 52-67.

[3] Barr, Nicholas, ed. 2001. Economic Theory and the Welfare State. The International Library of Critical Writings in Economics #132. Cheltenham, UK: Edward Elgar Publishing Limited.

[4] Barro, R., 1990. Government spending in a simple model of endogenous growth. *Journal of Political Economy* 98, 103-125.

[5] Becker, G., Murphy, K., 1988. The family and the state. *Journal of Law and Economics* 31, 1-18.

[6] Bishnu, Monisankar, 2011. Linking consumption externalities with optimal accumulation of human and physical capital and intergenerational transfers, mimeo, Indian Statistical Institute, New Delhi; *Journal of Economic Theory,* forthcoming

[7] Boldrin, M., Rustichini, A., 2000. Political equilibria with social security. *Review of Economic Dynamics* 3, 41-78.

[8] Bommier, A., Lee, R., Miller, T. and Zuber, S. 2010. Who Wins and Who Loses? Public Transfer Accounts for US Generations Born 1850 to 2090. *Population and Development Review*, 36, 1-26. doi: 10.1111/j.1728-4457.2010.00315.x

[9] Bossi Luca and Gulcin Gumus, 2012. Income inequality, mobility, and the welfare state: a political economy model. *Macroeconomic Dynamics*, forthcoming

[10] Bovenberg A. Lans 2010 New Social Risks, the Life Course, and Social Policy in in *Pension reform in Southeastern Europe: linking to labor and financial market reforms*, Robert Holzmann, Landis MacKellar and Jana Repansěk, editors.

[11] Brooks, David, 2012 The Structural Revolution, May 7, *New York Times*

[12] Conde Ruiz, J.I., Galasso, V., 2003. Early retirement. *Review of Economic Dynamics* 6, 12-36.

[13] Cooley, T.F., Soares, J., 1999. A positive theory of social security based on reputation. *Journal of Political Economy* 107, 135-160.

[14] Danish Economic Council, 2012, The Danish Economy - Spring 2012 (Dansk Økonomi - Foråret 2012), Copenhagen

[15] Del Rey, Elena and Lopez-Garcia, Miguel Angel, 2010. On welfare criteria and optimality in an endogenous growth model. CORE Discussion Papers 2010024, Université catholique de Louvain.

[16] Docquier, F., Paddison, O. and Pestieau, P. 2007. Optimal Accumulation in an Endogenous Growth Setting with Human Capital. *Journal of Economic Theory* 134, 361-378.

[17] Esping-Andersen, Gøsta. 1990. The Three Worlds of Welfare Capitalism. Princeton, NJ: Princeton University Press.

[18] Feldstein, M., and J. Leibman 2002. Social Security in *Handbook of Public Economics*, edited by A. Auerbach and M. Feldstein, Vol 4, chapter 32, Elsevier.

[19] Folbre, Nancy 2012. Define 'Welfare State,' Please, May 14, *New York Times*

[20] Garfinkel, Irwin, Timothy Smeeding, and Lee Rainwater. 2010. *Wealth and Welfare States: Is America a Laggard or Leader?* Oxford, UK: Oxford University Press.

[21] Gonzalez-Eiras, Martín and Niepelt, Dirk 2008. The future of social security. *Journal of Monetary Economics*, 55(2), 197-218.

[22] Hassler, J., Rodríguez Mora, J.V., Storesletten, K., Zilibotti, F., 2003. The survival of the welfare state. *American Economic Review* 93, 87-112.

[23] Heijdra, Ben J., Jochen O. Mierau, Laurie S.M. Reijnders, 2010. The Tragedy of Annuitization CESIFO WORKING PAPER No. 3141

[24] Iturbe-Ormaetxe, Iñigo and Guadalupe Valera, 2012. Social security reform and the support for public education. *Journal of Population Economics*, 25(2), 609-634

[25] Kaganovich, M., and Zilcha, I., 1999. Education, Social Security and Growth, *Journal of Public Economics*, 71, 289-309.

[26] Kaganovich, M., and Zilcha, I., 2012, Pay-as-you-go or funded social security? A general equilibrium comparison. *Journal of Economic Dynamics and Control*, 36(4), 455-467 http://dx.doi.org/10.1016/j.jedc.2011.03.015

[27] Konrad, K.A., 1995. Social security and strategic inter-vivos transfers of social capital. *Journal of Population Economics* 8, 315-326.

[28] Kotlikoff, Laurence J., 1998. Privatizing U.S. Social Security: Some Possible Effects on Intergenerational Equity and the Economy, Federal Reserve Bank of St. Louis *Review,* March/April.

[29] Lancia, Francesco and Russo, Alessia, 2010. A Dynamic Politico-Economic Model of Intergenerational Contracts. http://mpra.ub.uni-muenchen.de/24795/

[30] Lindbeck, Assar and Persson, Mats. 2003. The Gains from Pension Reform. *Journal of Economic Literature* , 41(1), 74-112

[31] Lindert, Peter H. 2004. *Growing Public: Social Spending and Economic Growth since the Eighteenth Century.* New York: Cambridge University Press.

[32] Lee, R., and A. Mason, 2011a, Population Aging and the Generational Economy – A Global Perspective, Edwar Elgar (www.ntaccounts.org)

[33] Lee, R. and A. Mason 2011b, Introducing age into national accounts, Chapter 7 in Lee and Mason (2011a).

[34] Miller, Tim 2011. The Rise of the Intergenerational State: Aging and Development in *Population Aging And The Generational Economy: A Global Perspective*, eds Ronald Lee and Andrew Mason, Edwar Elgar

[35] Mulligan, C. B. and Sala-i-Martin, X. 2004. Political and Economic Forces Sustaining Social Security, *Advances in Economic Analysis & Policy*: 4(1), Article 5.

[36] Norwegian Ministry of Finance, 2011, Nationalbudgettet 2012, Oslo Iin norwegian)

[37] Naito, Katsuyuki. 2012. Two-sided intergenerational transfer policy and economic development: A politico-economic approach, *Journal of Economic Dynamics and Control*, 10.1016/j.jedc.2012.02.008.

[38] OECD 2011. Education at a Glance 2011: OECD Indicators, OECD Publishing. http://dx.doi.org/10.1787/eag-2011-en

[39] Pecchenino, R.A, and P. S. Pollard, 1997. The Effects of Annuities, Bequests, and Aging in an Overlapping Generations Model of Endogenous Growth. *Economic Journal*, 107(1), 26-46.

[40] Petterson T., T.Petterson and A Westerberg, 2006, Generationsanalyser - omfördeling mellan generationer i en växande välfärdsstat. Expertgruppen för Studier i Samhällsekonomi (ESS) 2006:6.

[41] Poutvara, P. 2003. Gerontocracy Revisited: Unilateral Transfer to the Young May Benefit the Middle-aged. *Journal of Public Economics* 88, 161-174.

[42] Poutvaara, Panu. 2006. On the political economy of social security and public education. *Journal of Population Economics* 19(2), 345–365.

[43] Pogue, T.F., Sgontz, L.G., 1977. Social security and investment in human capital. *National Tax Journal* 30, 157-169.

[44] Rangel, A., 2003. Forward and backward intergenerational goods: Why is social security good for the environment? *American Economic Review* 93, 813-834.

[45] Razin, Assaf, Efraim Sadka and Phillip Swagel. 2002. "The Aging Population And The Size Of The Welfare State," *Journal of Political Economy*, 110(4), 900-918.

[46] Samuelson, Robert. 2011 The welfare state's reckoning *The Washington Post*, December 4

[47] Swedish Ministry of Finance, 2011, Vårpropositionen, Stockholm. (In swedish)

[48] Tanzi, V. and Schuknecht, L. 2000. *Public Spending in the 20th Century*. Cambridge University Press.

[49] Wang, Min. 2011. Optimal Education Policies under Endogenous Borrowing Constraints, *mimeo* CCER, China; forthcoming *Economic Theory*