



Institute for Empirical Research in Economics
University of Zurich

Working Paper Series
ISSN 1424-0459

Working Paper No. 236

**Fostering Within-Family Human Capital Investment:
An Intragenerational Insurance Perspective of Social Security**

Martin Barbie, Marcus Hagedorn and Ashok Kaul

March 2005

Fostering Within-Family Human Capital Investment: An Intragenerational Insurance Perspective of Social Security

Martin Barbie*, Marcus Hagedorn[†] and Ashok Kaul[‡]

This Version: March 2, 2005

Abstract

We propose an extended PAYG social security system that conditions pension benefits on the aggregate wage sum and on the wage of one's children. The latter increases parents' incentives to provide their children with good within-family education. However, since wages depend stochastically on parents' unobservable investment in their children's human capital, some insurance against the productivity risk of one's children is provided because retirement income still depends on aggregate wages. We analyze the effects of such a social security system on the endogenous distribution of human capital and compare it to real world systems which typically do not condition benefits on the wages of one's children. Our approach suggests a novel role for a well-designed social security system: it can foster human capital accumulation and act as an intra-generational insurance against productivity risk.

JEL classification: J24, H55, D61

Keywords: Human Capital Formation, Social Security, Intragenerational Insurance, Heterogenous Households

*University of Karlsruhe. barbie@wior.uni-karlsruhe.de

[†]University of Frankfurt and IZA. Hagedorn gratefully acknowledges support from the Deutsche Forschungsgemeinschaft (DFG). hagedorn@wiwi.uni-frankfurt.de

[‡]Kaul (correspondence): Gutenberg University Mainz, Welderweg 4, 55099 Mainz, Germany; and IZA, Bonn, Germany; email: ashok.kaul@uni-mainz.de.

The paper benefitted from comments by participants at the BGSE Workshop 2001 in Bonn, the ESPE 2001 in Athens, the Workshop on Intergenerational Risk Sharing 2002 in Fiesole, the EEA Meeting 2002 in Lausanne and the German Economic Association Annual Meeting 2002 in Magdeburg. The paper was revised while Barbie and Kaul were visiting the University of Zurich in the Summer of 2004. We would like to thank the Department of Economics for its hospitality, in particular Christian Ewerhart for inviting us.

1 Introduction

Retirement is organized differently in industrialized and rural economies. In the latter parents' old age consumption depends on what their children produce. In contrast, industrialized countries have developed public pension systems. Retirees get a share of aggregate wages and thus do not depend on their offspring's productivity. Both systems have their problems. Countries without a social security system expose the old generation to too much risk since consumption during old age strongly depends on one's children's productivity. In countries with social security, retirement income depends on the average productivity of the economy but is independent of one's offspring's productivity. Therefore, parents' incentives to invest in their children's education are low.¹ In a nutshell, rural economies provide no intra-generational insurance against the human capital risk of one's children whereas social security in industrialized economies provides no incentives to invest in one's children's human capital formation.

In this paper, we develop a highly stylized reform proposal for social security that combines the benefits of both systems. Retirement income depends both on aggregate wages and on the wages of one's children. This simultaneously provides parents with some insurance and with incentives to make sure that their children get a good education (to earn high wages). The kind of education we have in mind is done within the family. The empirical literature supports our view that the family plays an outstanding role in fostering skills, in particular when the children are young [e.g. Heckman (1999)]. However, it is important to notice that the outcome of parents' investment is stochastic, i.e. a high investment does not guarantee a high productivity. This is where our proposal comes into play. Retirement income is still linked to aggregate wages and therefore parents are partially insured against the income risk of their child. However, a fraction of the retirement income depends on one's children's wages in order to improve the low incentives of current social security systems to invest in one's children's human capital formation. Real world social security systems provide parents with full insurance against the productivity risk of their children and thus erode investment incentives in the education of their children. Other institutions have to step in and have to contribute to the formation of human capital, first of all public education. A conclusion one may draw is that the social security systems we observe in practice only go together with public education. In other words, one may argue that industrialized countries have outsourced the education of their children from the family to public education systems. But since education at home and at school are not perfect substitutes, this is presumably not optimal from a human capital perspective.²

¹See for example Cigno, Luporini and Pettini (2003) for further arguments confirming this.

²Boldrin and Montes (2002) consider the interaction of public education and social security. Public

Our paper starts from the observation that there is an insurance-incentive tradeoff in organizing retirement. The insurance aspect is intra-generational (parents of one generation are insured against low wages of their children) and incentives are inter-generational (investment in the next generation's human capital). Real world social security systems can be interpreted as a corner solution in this insurance-incentive tradeoff since they provide full insurance and no incentives to invest. Our model allows us to gain some new insights about an optimal social security benefit formula in a dynamic context (with endogenous human capital distribution). Furthermore, it sheds some light on popular proposals in European countries to link social security benefits to the number of children in order to provide incentives to have more children and thus weaken the demographic problems social security faces. It becomes clear that what matters for the viability of social security systems is not the size of the population, but the wage sum that is earned by the working population. Unemployed or low income/productivity individuals are thus not of much help in curing the social security crisis. We want to stress that better incentives for investment in education *within* the social security system may be a policy measure to achieve a more favorable long run distribution of human capital endowments and provide appropriate insurance against human capital risk of one's children. Our analysis suggests that parents' social security benefits should be conditioned (among other individual characteristics) on the contributions of one's children in order to establish optimal intra-generational risk sharing.

Most economists considering insurance aspects of social security have stressed the issue of intergenerational risk sharing.³ In a representative consumer framework, this is the obvious kind of risk management to consider. In such a framework, social security can act as insurance against factor share risk [Merton (1983)], as insurance against the state in which individuals are born [Gordon and Varian (1988), Bohn (1998)] and as insurance against demographic risk [Smith (1982), Demange and Laroque (1999), Bohn (2000)]. Most of the discussion on social security reform has also focused on the role of social security in providing intergenerational risk sharing [see Shiller (1999) for an excellent discussion]. There is only one paper, by Robert Shiller (1999), mentioning that social security can also have important functions in the pooling of risks among individuals in a single generation, i.e. the role for social security as intra-generational insurance. Shiller (1999) discusses the role of social Security as insurance against income risk in a heterogeneous population. In spite of practical limitations, however, this kind of risk sharing can at least partly be done by individual themselves by investing in a diversified

education, financed by middle age people, provides human capital for credit constrained young people. Pay back time is via a social security tax during middle age.

³The early literature on intergenerational risk sharing and social security includes Enders and Lapan (1982), Smith (1982) and Gordon and Varian (1988). More recent work includes Bohn (1998) and Abel (1999).

portfolio. Furthermore, the government already provides insurance against income risk by a progressive income tax system. In a related contribution, Cigno, Luporini and Pettini (2003) model the principal-agent problem of a government (principal) and parents (agent) where the latter exert unobservable effort to (stochastically) improve the well-being of their child. Their focus is, however, not on intra-generational issues.

The paper is organized as follows. Section 2 introduces our stochastic OLG economy with heterogeneous households. We also propose a social security system in which benefits are parameterized in a weight that determines how much of the benefits depends on the average wage and how much depends on the wage of one's child. We interpret this parameter as a public-pension incentive parameter. In section 3 we prove that under weak assumptions a unique equilibrium exists in the static household problem. We derive some comparative static results in section 4 and discuss the tradeoff between insurance and incentives in our general equilibrium framework. We identify four channels through which an adjustment in the social security incentive parameter affects households' investment decisions: an incentive effect, a redistributive effect, a general equilibrium effect and an insurance effect. We show that the sign of the overall effect depends on the income position of the household. We also provide some results concerning aggregate investment and welfare. Our main simplifying assumptions are discussed in Section 5. Section 6 concludes. An appendix contains all proofs as well as a sketch of a dynamic analysis of the human capital distribution and also illustrates the assumptions of our model and their implications of the model through an example.

2 The Model

The Basic Setup. Consider the following OLG model. Assume there are two generations alive at each period of time, old and young, and that there are $i = 1, \dots, N$ households in every generation. Every household has exactly one child at the end of the first period and lives for two periods. Time extends from zero to infinity ($t = 0, 1, \dots$). Saving is possible from first period net income only through investment in education e_i of one's child. Old age consumption is financed by social security as described below.

Uncertainty. Each household i , investing in education, is exposed to an idiosyncratic individual shock $\omega_i \in Y_i$ ($i = 1, \dots, N$) (genetic inheritance or just pure luck). Let $\Omega \equiv Y_1 \times \dots \times Y_N$ be a product of finite sets. We assume that shocks are independent and identically distributed over time according to probability $P(\omega_1, \dots, \omega_N) > 0$ for all $(\omega_1, \dots, \omega_N) \in Y_1 \times \dots \times Y_N$. This stochastic component is unknown to the household when investing in education.

Human capital The investment in education e_i in a child affects a child's human capital

endowment h_i according to a function $h_i = h(\cdot)$, where $h : \mathbb{R}_+ \times Y_i \rightarrow \mathbb{R}_+$. This function $h(e, \omega_{t,i})$ maps the amount e of the single consumption good invested in human capital and the individual shock $\omega_{t,i}$ into human capital (or equivalently in our model: efficiency units of labor) with $h(0, \omega_{t,i}) \geq 0$. Therefore, this function is a random variable that is parametrized by investment in education. In addition, this function is strictly increasing and strictly concave and twice continuously differentiable in its first argument. It also satisfies $\lim_{e \rightarrow 0} h_1(e, \omega_{t,i}) \rightarrow \infty$, where $h_1(\cdot)$ denotes the partial derivative with respect to e . In the following we will assume (dropping indices) that $h(e, \omega) = h(e) + \epsilon(\omega)$, where $\epsilon(\omega)$ is a random variable with zero mean. This can easily be generalized. Aggregate human capital is $H := \sum h_i$. Investment in education takes place in childhood within a family, for example in the form of time spent with the child and is thus not observable. Fostering this kind of investment through direct subsidies is thus ruled out.

Production. The production function is assumed to be linear: $F(H) = w \cdot H$, where H are efficiency units of labor (human capital) and $w > 0$ is the wage of unskilled labor.

Labor Market and Firms. Households inelastically supply efficiency units of labor on a competitive labor market. When young, a household receives gross wage $wh_{t,i}$, which depends on the unskilled wage w and its skill level $h_{t,i}$. The wage sum in the economy is $W := wH = w \sum h_i$

Social Security and Savings. There is a mandatory pay-as-you-go social security system. From the gross wage $wh_{t,i}$, a fraction τ is collected as social security payroll tax. The household i born at time t can spend the net income when young on consumption $c_{t,i}^y$ or on investment in education of its only child $e_{t,i}$. We consider a social security benefit function that consists of two parts. The first part is a fraction $1 - \alpha$ of an equal share of the wage sum in the economy, W/N . The second part is a fraction α of the wage of the household's own child, $wh_{t+1,i}$. The first part is exogenous for the household because the wage sum is perceived as fixed for the individual household. The second part can be influenced by the household through the investment decision.⁴

For $\alpha = 0$ the social security benefit only depends on the wage sum and is independent of the individual investment decision. This system insures the individuals completely against the human capital risk of their children. Real world social security systems typically offer this kind of full insurance. For $\alpha = 1$ the social security system conditions solely on the child's wage, but not on the wage sum. This system induces full incentives for investment in education, but supplies the individual with no insurance against human capital risk. This kind of system is a pure within-family transfer system. In reality, this system still exists in rural areas of less developed countries.

Household Income. The household income can be summarized as follows. When young,

⁴Kolmar (1997) suggests a similar scheme as a kind of child allowance in a deterministic model.

a household receives a net wage income of $V_i := wh_{t-1,i}(1 - \tau)$. When old, a household receives a two part social security benefit.

Household Preferences. Household preferences are represented by the utility function $u(c_{t,i}^y) + Ev(c_{t,i}^o)$, where $c_{t,i}^y$ denotes consumption when young and $c_{t,i}^o$ consumption when old. u and v are twice continuously differentiable, strictly increasing and strictly concave. Furthermore, we assume that $\lim_{b \rightarrow 0} u'(b) \rightarrow \infty$ and $v''' \geq 0$.

The Household's Problem. The individual has to decide how much to invest in the education of its child e_i . The individual's problem can be written as follows:

$$\begin{aligned} \max \quad & u(c_{t,i}^y) + Ev(c_{t,i}^o) \\ \text{s.t.} \quad & c_{t,i}^y + e_{t,i} = w \cdot h(e_{t-1,i}, \omega_{t,i}) \cdot (1 - \tau), \\ & c_{t,i}^o = \tau \cdot \left[(1 - \alpha) \cdot \frac{W}{N} + \alpha \cdot w \cdot h(e_{t,i}, \omega_{t+1,i}) \right] \end{aligned}$$

where $0 < \alpha \leq 1, 0 < \tau < 1$. The problem can be transformed by substituting the expressions for c^y and c^o into the objective function. Furthermore, individuals take prices (wages w and interest rates R) and repayments from the wage sum W as given, where W will in equilibrium be equal to $w \cdot \sum_{j=1}^N h(e_{t,j}, \omega_{t+1,j})$. In addition, we use the definition $V := w \cdot h_{t-1,i} \cdot (1 - \tau)$. We skip indices and ω 's and slightly abuse notation to obtain the following problem for a period t born individual i :

$$\max_{0 \leq e \leq V} u(V - e) + Ev \left(\tau \cdot \left[(1 - \alpha) \cdot \frac{1}{N} \cdot W + \alpha \cdot w \cdot h(e, \cdot) \right] \right). \quad (1)$$

In summary, the individual when young invests part of his initial endowment V in the education e of his child. The payoff of this investment is received during old age. It consists of a fraction of the wage sum and a fraction of the wage of his child.

3 Existence and Uniqueness of Equilibrium

The first-order conditions (FOC) for the individual problem (using the boundary behavior and the strict concavity of the utility functions u and h) characterize a perfect foresight competitive equilibrium. We have $\forall i = 1, \dots, N$:

$$u'(V_i - e_i) = E\{v'[\tau w \cdot (\alpha h(e_i, \omega_i) + (1 - \alpha) \frac{1}{N} \sum_{j=1}^N h(e_j, \omega_j))] \cdot \tau \alpha w \cdot h_1(e_i, \omega_i)\}. \quad (2)$$

To simplify notation we will refer to the FOC as follows:

$$u'(V_i - e_i) = RHS(e_i, \alpha).$$

We obtain the following quite general result:

Proposition 1 *For any given profile of positive after tax incomes (V_1, \dots, V_N) a solution to the static household problem exists and is unique.*

4 Comparative Statics

In this section, we analyze how changes in the social-security incentive parameter α affect households' consumption and investment decisions e_i . As benchmarks, we first consider a case where full insurance provision is optimal because moral hazard is absent and a case where household heterogeneity is ignored. We then turn to the general case and describe how households' consumption and investment decisions depend on the distribution of wealth V_i for a given α , i.e. for a given insurance-incentive tradeoff. Finally, we turn to the comparative statics with respect to α , i.e. the question how *changes* in the insurance-incentive tradeoff affect the distribution of investment as well as the aggregate wealth level and welfare.

Benchmark 1: Optimality of Full Insurance

Consider first the benchmark case where e_i is observable and contractible and the government can ensure that all households invest the same amount e , $e_i = e \forall i$. This benchmark abstracts from incentive problems due to moral hazard and from heterogeneity in income. In particular, differences in the endowments of households and the corresponding differences in the level of redistribution built into the social security benefit formula do not feed back on investment e . The benchmark therefore isolates the intragenerational insurance aspect of social security. We have the following result:

Proposition 2 *Without incentive problems and abstracting from the effects of heterogeneity on investment, households benefit from better insurance. In this case full intragenerational insurance, $\alpha = 0$, is optimal.*

This result only holds for the *isolated* insurance effect. $\alpha = 0$ can never be optimal in the model. The reason is that $\alpha = 0$ would imply $e_i = 0$ for all households i , because under full insurance no investment incentive exists due to the possibility of free riding. We have avoided this in the benchmark by separating insurance from incentive aspects through fixing $e_i = e$ exogenously. In real world social security systems we observe $\alpha = 0$.

But this only means that within family investment in education is inefficiently low. Other institutions also contribute to the formation of human capital, first of all public education. One conclusion we may draw is that the social security systems we observe in practice only go together with public education. In other words industrialized countries have outsourced the education of their children.

Benchmark 2: Optimal Insurance for Homogenous Households

Now we turn back to the case with an incentive problem. What does the optimal α look like? To get some insight into this problem for the case homogenous households (i.e. $V_i = V$ for all i), we maximize the utility of the representative consumer over the set of competitive equilibria associated with different values of α . This procedure is the same one as used in theory of optimal (Ramsey) taxation to find the optimal tax rates. Since competitive equilibria are characterized by (2), the problem becomes:

$$\begin{aligned} \max_{e, \alpha} \quad & u(V - e) + Ev(\tau \cdot [(1 - \alpha)wh(e) + \alpha w(h(e) + \epsilon)]) \\ \text{s.t.} \quad & u'(V - e) = E[v'(\tau \cdot [(1 - \alpha)wh(e) + \alpha w(h(e) + \epsilon)]) \cdot \tau \alpha wh'(e)]. \end{aligned}$$

In the presence of the incentive problem, can it become optimal to set $\alpha = 1$, i.e. to have no social security system at all? Can the incentive problem even become so severe that it is indeed optimal to choose $\alpha > 1$, i.e. people have to pay a lump sum tax in old age if they have no (productive) children at all? If the incentive problem becomes sufficiently severe, this it might seem plausible that such a choice of α may become optimal. we show that this is can never happen.

Proposition 3 *If households are homogenous, the optimal α is strictly larger than 0 and strictly less than 1.*

The proposition shows that in the case of homogenous households, there is always some scope for a pay-as-you-go social security system. Note that the assumption $v''' > 0$ is used in the derivation of the result. Intuitively, the presence of precautionary saving (due to $v''' > 0$) and $\alpha \geq 1$ will lead to a too high choice of e . To see this, note that a social planner who can choose both α and e will set $\alpha = 0$ and e according to:

$$u'(V - e) = E[v'(\tau \cdot wh(e)) \cdot \tau wh'(e)].$$

On the other hand, the choice of the consumer is determined by:

$$u'(V - e) = E[v'(\tau \cdot [wh(e) + \alpha w\epsilon]) \cdot \tau \alpha wh'(e)].$$

When $\alpha \geq 1$, the consumer will choose a higher e because the right hand side of his first order condition is for a fixed e larger than the right hand side of the first order condition of the social planner. This comes from the direct effect of α on the marginal productivity from the consumer's perspective of investment in education, $\tau\alpha wh'(e)$, and also from the increase in riskiness in the second period, $\alpha w\epsilon$, which leads to additional precautionary saving. Thus lowering α brings the e closer to its optimal level and makes the consumer carry less risk, which improves his utility.

Heterogeneous Households

The Distribution of Investment for given Incentives

Now consider how the distribution of wealth affects the investment decisions of households for any given redistributive social security system (parameterized in $0 < \alpha < 1$).

Proposition 4 (a) *The rich invest more: If $V_1 < \dots < V_N$, then $e_1 < \dots < e_N$.*

(b) *The rich consume more: If $V_1 < \dots < V_N$ then $c_1^y = V_1 - e_1 < \dots < V_N - e_N = c_N^y$.*

These two results show that the distribution of wealth is mapped one-to-one to a distribution of investment and consumption for any given level of redistribution. This has to be kept in mind when changing the insurance-incentive tradeoff within a social security system.

The Effects of Higher Incentives on Household Decisions

Now we disentangle the channels through which changes in incentive provision within a social security system interact with insurance provision, redistribution and how this feeds back to the distribution of investment. This is achieved by considering how changes in α affect the expected marginal revenue of investment of the households which is given by $RHS(\cdot)$. We know from the FOC that investment e_i rises if $RHS(\cdot)$ rises.

$$\frac{\partial RHS}{\partial \alpha} = E\left\{v''\tau\left[-\frac{W(\alpha)}{N} + w[h(e) + \epsilon] + (1 - \alpha)\frac{W'(\alpha)}{N}\right]\right\}\alpha wh'(e)\tau + E[v']wh'(e)\tau \quad (3)$$

The marginal revenue changes due to two effects. The first summand is an indirect effect, the second summand a direct incentive effect. The sign of the indirect effect depends on the term in braces. The indirect effect accounts for three changes due to a change in α . First, a change in α entails changes in the redistributive properties of our social security system (redistributive effect). Second, changes in α will feed back through changes in investment incentives on the realization of the aggregate wage sum $W(\alpha)$ (general equilibrium effect). Third, the insurance arrangement of our social security

system is affected (insurance effect). We will neglect for a moment the general equilibrium effect of a change in α , $W'(\alpha)$.

Incentive Effect. The direct effect is positive since a higher incentive parameter raises the individual's incentives to invest for a given expected wealth in the second period.

Redistributive Effect. The redistributive effect is described by the term $-W(\alpha)/N + w \cdot h(e)$. Its sign depends on the income position of the household. If the household income is below the average income, $W(\alpha)/N > w \cdot h(e)$, this effect is also positive. It is negative for above average income households, since a higher α implies less redistribution via the social security system. This benefits the high income households whose marginal revenue in old age thus falls due to higher after social security tax income. A similar argument shows that a higher α hurts the poor. Regarding investment incentives, this implies less investment by the rich and more investment by the poor.

Insurance Effect. The insurance effect is described by the term $E[v''(\cdot) \cdot \tau \cdot w \cdot \epsilon]$. Under our assumptions $E[\epsilon] = 0$ and $v''' > 0$ this term can be shown to be positive. A proof and a detailed discussion of this effect will be given in the appendix in the proof of proposition 2. The interpretation is as follows: a rise in α reduces the insurance coverage of the households (recall that $\alpha = 0$ is the full insurance case) and this in turn lowers the second period expected utility, thus raising expected marginal utility in period 2 as described by $RHS(\cdot)$. As a consequence, the insurance effect of a higher α will raise investment.

Thus, neglecting the general equilibrium effect, a higher incentive parameter has an unambiguously positive effect on the investment of the poor, but an ambiguous effect on the rich. Since the effect of a change in α on investment behavior is generally ambiguous, we give a sufficient condition under which more can be said about the changes in the investment behavior of the households. For comparative statics experiments that we conduct, it has proved to be useful to make assumption guaranteeing some kind of monotonicity (supermodularity, for example, see Topkis (1998) and Gollier (2001), section 7). We impose an assumption that resembles supermodularity in the sense that a cross derivative is positive.

Assumption 1 $\frac{\partial}{\partial e} \frac{\partial RHS(e, \alpha)}{\partial \alpha} < 0$

Assumption 1 says that the changes - due to higher incentives - in the marginal revenue of investment should fall monotonically in the level of investment e . This assumption is consistent with our previous finding, namely that the redistributive effect in isolation implies $\partial RHS/\partial \alpha > 0$ for small e , i.e. for poor households, and $\partial RHS/\partial \alpha < 0$ for large e , i.e. for rich households. It strengthens this finding by imposing that the overall effect of this fall in $\partial RHS/\partial \alpha$ should be monotonic in e . This assumption is only sufficient but not

necessary to obtain the following proposition. This can be seen in a illustrative example in the appendix. The assumption is satisfied, for example, for the case of log-utility.

Proposition 5 *If assumption 1 holds then either (a) or (b) is true:*

- (a) *Higher incentives (higher α) induce higher investment e_i for all households i .*
- (b) *If there exists a household j for which investment e_j falls although the incentives rise, then e_k falls for all $k > j$. Thus household j is a cutoff household.*

The proposition shows that higher incentives to invest need not lead to higher investment over the whole range of households. In particular, those households who are rich gain by the higher incentives through the redistributive effect and therefore might reduce their investment. The reason for this is, as pointed out above, that in our model higher incentives imply less redistribution, which makes the rich better off in expected utility terms. This lowers their expected marginal utility of second period consumption and consequently lowers investment.

In summary, we have identified four channels through which a change in the social-security incentive parameter affects households' investment decisions, namely an incentive effect, a redistributive effect, a general equilibrium effect and an insurance effect. We were able to sign the effects and derived some results about the overall effect. In particular, the sign of the overall effect depends on the income position of the household. For below average income households higher investment incentives will indeed induce higher investment. For above average income households this need not be true.

Aggregate Effects of Higher Incentives to Invest

We now turn to the aggregate effects of a policy that sets higher incentives to invest in education by increasing α . As we saw in Proposition 5, individually the effect on household investment is ambiguous. Under mild additional assumptions, we obtain a clear-cut result on the aggregate level. Furthermore, we obtain an analogue to Proposition 3 for the heterogeneous household case.

Proposition 6 (a) *Suppose that Proposition 5 holds and that the richest household has a relative risk aversion smaller than or equal to one.*

Then aggregate investment increases in α : $W'(\alpha) > 0$.

(b) *Suppose that the households with above-average income, i.e. household i with $wh(e_i(\alpha)) > \frac{W(\alpha)}{N}$, have a relative risk aversion smaller or equal to one.*

Then aggregate investment increases in α : $W'(\alpha) > 0$.

(c) *Suppose that Proposition 5 holds and that the richest household has a relative risk aversion smaller or equal to one. Suppose furthermore that the planner designing*

social security evaluates social welfare according to a Utilitarian welfare function. Then some strictly interior $\alpha \in (0, 1)$ maximizes welfare.

This proposition strengthens our finding so far by imposing mild additional assumptions. It justifies our social security reform proposal on two grounds. First, it makes a statement about aggregate investment although individually the effect of providing higher incentives to invest was ambiguous. Second, it trades off winners and losers of such a reform by imposing a specific but widespread welfare function. Our analysis was not meant to be quantitative. It was meant to be suggestive in the sense that we pointed out that current social security reform proposals tend to neglect a particular but important insurance-incentive tradeoff. Our punchline is that current social security systems resolve this tradeoff by imposing a corner-solution that is likely to be suboptimal under quite general assumptions.

5 Discussion of Simplifying Assumptions

5.1 Altruism

By no means we want to claim the absence of altruism. Our assumption is that altruism is not strong enough (from a social perspective) to provide sufficient incentives. In addition aggregate externalities in human capital accumulation itself or of human capital accumulation on crime rates (see Heckman(1999)) are not fully taken into account. For simplicity we choose to work within a standard OLG environment, what means we assume no altruism at all. Since our aim is not a serious quantitative exercise this assumption is not too restrictive for our purposes. Qualitatively all our results still hold if we allow for some altruism.

5.2 Capital Accumulation

A main simplifying assumption of our model is that saving is possible only through investment in education e_i of one's child. Physical capital accumulation is ruled out but this does not change our results. First, physical capital accumulation would decrease the dependence of parents on the success of their children. But this is a quantitative statement. A different α between zero and one may be optimal but our reform proposal is still reasonable. Second, one may argue that, because of capital accumulation, interest rates are endogenous and this affects our results. This is not the case. None of our results are quantitative and do not depend on the size of interest rates. They hold for all time preference rates and for all utility functions.

Allowing for capital accumulation would only be warranted if we were interested in the

interaction of physical and human capital accumulation on the hand and social security on the other hand. But this paper is about the interaction of (within family) education and retirement. Thus we think that our results are unaffected if physical accumulation was allowed for.

5.3 Endogenous Fertility

A main point of our paper is that human capital and not the sheer size of the next generation matters. To highlight this point, we allowed for only one child per person and focused on the offspring's amount of human capital h . In a more realistic model, where the number of children is endogenous, we think of human capital h as the sum of the human capital of all children (of one person). h is then the product of the number of children and their individual human capital. Instead of deciding about the number of children and the amount of human capital for each child, parents in our model only decide about the sum of the human capital of all their children. Again we are not claiming that this is realistic but that this assumption makes the model simpler and does not change our results.⁵

5.4 Linear Incentives

We started from the observation that there is an incentive-insurance tradeoff in organizing retirement. Instead of deriving the optimal contract in a principal(government)-agent(parents) framework we restricted ourselves to a simple linear contract. Our contract only depends on aggregate wages and the wage of one's own child. If the Law of Large Numbers applies, the same arguments as in Holmström (1979,1982) and Mookherjee (1984) show that this is not restrictive. Nevertheless the optimal contract would be more sophisticated and highly nonlinear. The nonlinearity crucially depends on the stochastic environment and we have to make strong assumptions, such as the monotone likelihood ratio condition, to derive even elementary properties (monotonicity for example). All characterizations would depend on the government exactly knowing the stochastic environment. Our simple linear scheme is robust to these problems and thus more accessible to politicians. We think that not only a linear scheme is sufficient to illustrate our idea but is also a reasonable benchmark; in the same way that (linear) Ramsey taxation is a useful benchmark for an information theoretic approach to taxation, as propagated by Mirrlees (1974).

⁵Schweizer (1996) points out the technical difficulties of modelling quantity and quality of children simultaneously.

6 Conclusion

We have developed a general equilibrium stochastic OLG model with intra-generational heterogeneity in human capital and endogenous human capital distribution. In this framework, we analyzed a novel insurance aspect of social security. We suggested that a well designed social security system can act as insurance against the human capital risk of one's children. An optimal social security scheme, however, has to tradeoff insurance provision against incentives for appropriate investment in education. This insurance-incentive aspect was built into the benefit formula of a proposed social security scheme. Our analysis allowed us to interpret real-world social security systems as corner solutions in this insurance-incentive tradeoff.

We think our analysis highlights an aspect which was neglected in the literature. Better incentives for investment in education *within* the social security system may be one policy measure to achieve a more favorable long run distribution of human capital endowments and provide appropriate insurance against human capital risk of one's children. It is superior to a proposal that suggests an increase in the number of children. What matters is the amount of human capital of each generation.

Furthermore, our analysis highlights that redesigning social security along the insurance-incentive dimension will inevitably have distributional consequences that will tend to feed back on the incentives to invest in education, both on the individual level as well as on the aggregate level. In the appendix we provide a sketch how such a dynamic analysis could be carried out within the framework of our paper. A detailed analysis is, however, beyond the scope of this paper.

Appendix

Proof of Proposition 1

The following assumption is a sufficient condition for the existence of a unique fixed point of the first order equation system. It is e.g. satisfied if shocks are additive as was assumed in the main text.

Assumption 2 For $\hat{e} \neq e$, we have $\forall \omega, \omega' \in Y_1 \times \dots \times Y_N$:

$$\sum_{j=1}^N h(e_j, \omega_j) \geq \sum_{j=1}^N h(\hat{e}_j, \omega_j) \iff \sum_{j=1}^N h(e_j, \omega'_j) \geq \sum_{j=1}^N h(\hat{e}_j, \omega'_j).$$

This assumption compares aggregate human capital as a result of two different education profiles. It imposes the following: if in one state aggregate human capital resulting from a certain investment profile is higher than that resulting from another investment profile, then this should hold uniformly in all states.

Uniqueness. Suppose there are two or more fixed points. Then three cases are possible:

- $0 < e_i \leq \widehat{e}_i \quad \forall i = 1, \dots, N$ with strict inequality for some i .
From the FOC (2) the left-hand side of all equations does not decrease by replacing e with \widehat{e} . Using the strict concavity of $h(\cdot, \omega_i)$, for individual i the right-hand side is decreasing. To see this, note first, that $w \cdot \frac{1}{N} \sum_{j=1}^N h(e_j, \omega_j)$ is increasing. This fact together with the concavity of v and h imply the claim. Thus no other fixed point with this property can exist.
- $e_i \geq \widehat{e}_i > 0 \quad \forall i = 1, \dots, N$ with strict inequality for some i .
The same argument as before applies with reversed inequalities.
- $e_i < \widehat{e}_i$ for some i and $e_j > \widehat{e}_j$ for some j .
Two further subcases have to be distinguished. First assume that $\sum_{j=1}^N h(e_j, \omega_j) \leq \sum_{j=1}^N h(\widehat{e}_j, \omega_j)$. By assumption 2, this expression holds for all $(\omega_1, \dots, \omega_N)$ if it holds for one such tuple. Consider an individual i for which e_j increases. Then the left-hand side in (2) increases, while $w \cdot \alpha \cdot h(e_i, \omega_i)$ increases and also $w \cdot \frac{1}{N} \sum_{j=1}^N h(e_j, \omega_j)$ increases. Thus $v'(\cdot)$ decreases and due to the strict concavity of $h(\cdot, \omega_i)$ the expression outside $v'(\cdot)$ decreases too. The reverse subcase is handled similarly.

Existence. Existence of a fixed point for the equation system (2) will be proved by Brouwer's fixed point theorem. Consider a cleaned up FOC for i :

$$u'(V_i - e_i) = E \left\{ v' \left(\tau \cdot \left((1 - \alpha) \frac{W}{N} + \alpha \cdot w \cdot h(e_i) \right) \right) \cdot \tau \cdot \alpha \cdot w \cdot h'(e_i) \right\}. \quad (4)$$

Given the boundary behavior of u, v and h , the solution to the individual optimization problem is interior, i.e. $0 < e_i < V_i$ (which also justifies the use of FOCs), since the given wage w is positive. This can be seen as follows. If e_i were equal to V_i , the left-hand side would be infinity while the right hand side would be a finite number. If e_i were equal to zero the right hand side would be infinity because of the Inada condition on h and the left hand side would be finite. Also, by the strict concavity of the problem, the solution will be unique for given W .

Now recall that W is a function depending on the shocks ω . By the maximum theorem the maximizer is a continuous function of W (where W is viewed as a vector in \mathbb{R}^L with L is the number of elements in Ω). Thus we have a continuous function $k_i : \mathbb{R}_+^L \rightarrow [0, V_i]$ for each individual i with $k_i(W) = \arg \max (1)$. To determine prices which are consistent with the choice of the individuals, i.e. in order to have a perfect foresight equilibrium, we introduce a map \widetilde{W} which assigns the wage sum corresponding to the choice of individuals. Viewed as a function from \mathbb{R}^N to \mathbb{R}^L , this map is continuous. Thus:

$$\widetilde{W} : \times_{i=1}^N [0, V_i] \rightarrow [\underline{F}, \overline{F}]^L$$

where $\overline{F} = \max_{\omega \in \Omega} w \cdot \sum_{i=1}^N h(V_i, \omega_i)$ and $\underline{F} = \min_{\omega \in \Omega} w \cdot \sum_{i=1}^N h(0, \omega_i)$.

To be able to obtain a fixed point, we thus consider a map from

$$[0, \overline{F}]^L \times_{i=1}^N [0, V_i]$$

into itself defined as

$$\Phi(W, e) = W(e) \times \times_{i=1}^N k_i(W, w, R).$$

So by Brouwer's fixed point theorem, there exists a fixed point.

Proof of Proposition 2

Under $e_i = e \forall i$ it follows that $h_i = h(e) + \varepsilon_i$ and thus $w_i = w \cdot h_i = wh(e) + w\varepsilon_i$ and $W = Nwh(e)$ from $\sum \varepsilon_i = 0$. We show that utility falls if α is raised for $\alpha \in [0, 1]$ in the case of exogenously fixed and identical investments e . Consider $\frac{\partial}{\partial \alpha} E[v(\cdot)]$. We have $\frac{\partial}{\partial \alpha} E[v[\tau((1-\alpha)W/N + \alpha wh(e) + \alpha w\varepsilon)]] = E\{v'(\tau W/N + \tau \alpha w\varepsilon) \cdot \tau \cdot [w(h(e) + \varepsilon) - W/N]\} = E\{v'(\tau W/N + \tau \alpha w\varepsilon) \cdot \tau w\varepsilon\}$. We will show that this term is negative. It suffices to show that $E[v'(c + d\varepsilon) \cdot \varepsilon] < 0$, where c, d are positive constants. We have

$$\begin{aligned} E[v'(c + d\varepsilon) \cdot \varepsilon] &= \sum_{\varepsilon > 0} v'(c + d\varepsilon)\varepsilon f(\varepsilon) + \sum_{\varepsilon < 0} v'(c + d\varepsilon)\varepsilon f(\varepsilon) \\ &< \sum_{\varepsilon > 0} v'(c)\varepsilon f(\varepsilon) + \sum_{\varepsilon < 0} v'(c)\varepsilon f(\varepsilon) = 0 \end{aligned}$$

The inequality follows from $v'' < 0$, the second equality used $E[\varepsilon] = 0$. This implies the claim. It follows with the same arguments as in this proof that $E[v''(\cdot) \cdot \tau \cdot \varepsilon] > 0$ under our assumptions $E[\varepsilon] = 0$ and $v''' > 0$. This was claimed above in the text.

Proof of Proposition 3

We form the Lagrangian

$$\mathcal{L} = u(V - e) + Ev(\tau \cdot [wh(e) + \alpha w\varepsilon]) + \lambda [u'(V - e) - E[v'(\tau \cdot [wh(e) + \alpha w\varepsilon]) \cdot \tau \alpha wh'(e)]],$$

where λ denotes the Lagrange multiplier.

The necessary first order conditions are

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial e} &= u'(V - e)(-1) + E[v'(\tau \cdot [wh(e) + \alpha w\varepsilon]) \cdot \tau wh'(e)] \\ &\quad + \lambda \left[\begin{array}{c} u''(V - e)(-1) - E[v''(\tau \cdot [wh(e) + \alpha w\varepsilon]) \cdot \tau^2 w^2 \alpha (h'(e))^2] \\ - E[v'(\tau \cdot [wh(e) + \alpha w\varepsilon]) \cdot \tau \alpha wh''(e)] \end{array} \right] \\ &\stackrel{!}{=} 0 \end{aligned} \tag{5}$$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \alpha} &= E[v'(\tau \cdot [wh(e) + \alpha w\varepsilon]) \cdot \tau w\varepsilon] - \lambda E[v'(\tau \cdot [wh(e) + \alpha w\varepsilon]) \cdot \tau \alpha wh'(e)] \\ &\quad - \lambda E[v''(\tau \cdot [wh(e) + \alpha w\varepsilon]) \cdot \tau^2 w^2 \varepsilon h'(e)] \stackrel{!}{=} 0 \end{aligned} \tag{6}$$

and

$$\frac{\partial \mathcal{L}}{\partial \lambda} = u'(V - e) - E[v'(\tau \cdot [wh(e) + \alpha w\varepsilon]) \cdot \tau \alpha wh'(e)] \stackrel{!}{=} 0. \tag{7}$$

For notational simplicity we suppress from now on the arguments in the functions u and v . From (5) and (7) we get:

$$(1 - \alpha) \cdot E[v'(\cdot) \cdot \tau wh'(e)] = \lambda \cdot \left[u''(\cdot) + E[v''(\cdot) \cdot \tau^2 w^2 \alpha (h'(e))^2] + E[v'(\cdot) \cdot \tau \alpha wh''(e)] \right]. \tag{8}$$

Rearranging (6), we obtain

$$E[v'(\cdot) \cdot \tau w \epsilon] = \lambda \cdot [E[v'(\cdot) \cdot \tau \alpha w h'(e)] + E[v''(\cdot) \cdot \tau^2 w^2 \epsilon h'(e)]] .$$

Further, note that

$$E[v'(\cdot) \cdot \tau w \epsilon] = E[v'(\cdot) \cdot \tau w] \cdot E[\epsilon] + Cov(v'(\cdot) \cdot \tau w, \epsilon) .$$

Therefore, with $E[\epsilon] = 0$,

$$Cov(v'(\cdot) \cdot \tau w, \epsilon) = \lambda \cdot [Cov(v''(\cdot) \cdot \tau^2 w^2 h'(e), \epsilon) + E[v'(\cdot) \cdot \tau \alpha w h'(e)]] .$$

Given that $v'' < 0$, we have $Cov(v'(\cdot) \cdot \tau w, \epsilon) < 0$. Further, since $v''' > 0$, it holds $Cov(v''(\cdot) \cdot \tau^2 w^2 h'(e), \epsilon) > 0$. Thus $\lambda < 0$, and therefore in (8), we must have $1 - \alpha > 0$, and thus $\alpha < 1$.

To see that $\alpha > 0$, note that with $\alpha = 0$, no investment in education is undertaken, i.e. $e = 0$. Using that $h'(0) = \infty$ it is not difficult to show that is never optimal to set $\alpha = 0$.

Proof of Proposition 4

Part a: Consider the FOC of the static problem, $u'(V_i - e_i) = RHS(e_i, \alpha)$, where we recall that $RHS(\cdot) := E[v'(\cdot)]\alpha w h'(e)\tau$ denotes the right hand side of the FOC as a function of α and e_i . Assume $V_i > V_j$ and $e_i \leq e_j$. It follows that $V_i - e_i > V_j - e_j$. Thus $u'(V_i - e_i) < u'(V_j - e_j)$ by concavity of $u(\cdot)$. Differentiating $RHS(\cdot)$ yields:

$$\frac{\partial RHS(\cdot)}{\partial e} = E[v''(\cdot)] \cdot [\alpha w h'(e)\tau]^2 + E[v'(\cdot)]\alpha w h''(e)\tau < 0$$

This gives $RHS(e_j, \alpha) < RHS(e_i, \alpha) = u'(V_i - e_i) < u'(V_j - e_j)$, where the first inequality follows from $\partial RHS/\partial e < 0$, the equality follows from the FOC for household i and the last inequality follows from the concavity of $u(\cdot)$. But this contradicts $RHS(e_j, \alpha) = u'(V_j - e_j)$ by the FOC of household j .

Part b: Assume $V_i > V_j$. By part (a), $e_i > e_j$. Thus $RHS(e_j, \alpha) > RHS(e_i, \alpha)$. By the FOC this gives $u'(V_j - e_j) > u'(V_i - e_i)$.

Proof of Proposition 5

We prove part b) of the proposition in four steps. Assume α is raised from α^1 to $\alpha^2 > \alpha^1$ and there exists some j such that $e_j^2 < e_j^1$, where e_j^z denotes the investment of household j under the incentive scheme α^z ($z = 1, 2$). Step 1. Claim: $0 < RHS(e_j^1, \alpha^1) - RHS(e_j^1, \alpha^2)$. Proof: We have $RHS(e_j^1, \alpha^1) = u'(V - e_j^1) > u'(V - e_j^2) = RHS(e_j^2, \alpha^2) > RHS(e_j^1, \alpha^2)$, where the first and last equality follow from the FOC, the first inequality follows from $e_j^1 > e_j^2$, the last inequality follows from $e_j^1 > e_j^2$ and $\partial RHS/\partial e < 0$. Thus it follows $RHS(e_j^1, \alpha^1) > RHS(e_j^1, \alpha^2)$, as claimed. Step 2. Claim: $RHS(e_j^1, \alpha^1) - RHS(e_j^1, \alpha^2) > RHS(e_j^2, \alpha^1) - RHS(e_j^2, \alpha^2) > 0$. Proof: We have

$$RHS(e_j^1, \alpha^1) - RHS(e_j^1, \alpha^2) = - \int_{\alpha^1}^{\alpha^2} \frac{\partial RHS(e_j^1, a)}{\partial a} da > - \int_{\alpha^1}^{\alpha^2} \frac{\partial RHS(e_j^2, a)}{\partial a} da = RHS(e_j^2, \alpha^1) - RHS(e_j^2, \alpha^2) .$$

The inequality follows from assumption 1 and $e_j^2 < e_j^1$. The left hand side is positive from step 1. The right hand side is positive by replacing e_j^1 with e_j^2 in step 1. Step 3. Claim: $RHS(e_k^1, \alpha^2) < RHS(e_k^1, \alpha^1)$ for $k > j$. Proof: We have

$0 < RHS(e_j^1, \alpha^1) - RHS(e_j^1, \alpha^2) < RHS(e_k^1, \alpha^1) - RHS(e_k^1, \alpha^2)$, where the first inequality follows from step 1 and the second inequality from $e_k^1 > e_j^1$ and step 2. Step 4. Claim: $e_k^1 < e_k^2$ implies $RHS(e_k^2, \alpha^2) < RHS(e_k^1, \alpha^2) < RHS(e_k^1, \alpha^1) = u'(V_k - e_k^1) < u'(V_k - e_k^2)$. Proof: The first inequality follows from $\partial RHS/\partial e < 0$, the second from step 3, the equality from the FOC and the last inequality from the concavity of $u(\cdot)$ and our assumption $e_k^1 < e_k^2$. But this is a contradiction. Hence $e_k^1 > e_k^2$, as claimed in part b) of the Proposition. Part a) is also a possible outcome of our model is shown in an example below.

Proof of Proposition 6

To prove part (a), recall that it suffices to show that the right hand side of the first order condition (3) increases in α . We use the estimate:

$$E\{v'' \cdot \left[\tau \left[-\alpha \frac{W}{N} + \alpha w[h + \epsilon] \right] \right\} + E[v'] \geq E\{v'' \cdot \left[\tau \left[(1 - \alpha) \frac{W}{N} + \alpha w[h + \epsilon] \right] \right\} + E[v'].$$

Thus if $-\frac{v''(x)x}{v'(x)} \leq 1$ for all x , $E\{v'' \left[\tau \left[(1 - \alpha) \frac{W}{N} + \alpha w[h + \epsilon] \right] \right\} + E[v'] \geq 0$, where $x := \tau \left[(1 - \alpha) \frac{W}{N} + \alpha w[h + \epsilon] \right]$ is used as in the right-hand-side of the above estimate.⁶ Thus for households with relative risk aversion less than or equal to 1 we now know the sign of all summands of $RHS'(\alpha)$ except for the term $E\{v''(\cdot) \cdot \tau \cdot [(1 - \alpha) \frac{W'(\alpha)}{N}]\}$.

Assume $W'(\alpha) \leq 0$. Then $E\{v''(\cdot) \cdot \tau \cdot [(1 - \alpha) \frac{W'(\alpha)}{N}]\}$ is also positive and in this case all households with risk-aversion smaller than one would invest more. In particular the richest household would invest more. But $W'(\alpha) \leq 0$ implies that there exists at least one household i who invests less. Under proposition 5, all households richer than i (surely including the richest household), also invest less, a contradiction.

Suppose the conditions in part (b) of proposition 6 hold. Suppose $W'(\alpha) \leq 0$. For all households with relative risk aversion less than or equal to 1, if $W'(\alpha) \leq 0$, investment increases. For all households with $\frac{W(\alpha)}{N} \geq w \cdot h(e)$, if $W'(\alpha) \leq 0$, investment also increases because $E\{v''(\cdot) \cdot \tau \cdot \left[-\frac{W(\alpha)}{N} + w \cdot h(e) + (1 - \alpha) \frac{W'(\alpha)}{N} \right]\} \geq 0$ and $E[v''(\cdot) \cdot \epsilon] > 0$. Thus, $W'(\alpha) > 0$ must hold and we obtain a contradiction.

For part (c), denote by $e_i(\alpha)$ the investment choice of household i for α . It is implicitly defined by the first order condition:

$$u'(V_i - e_i(\alpha)) = E \left[v' \left(\tau \cdot \left[(1 - \alpha) \frac{w}{I} \sum_j h(e_j(\alpha)) + \alpha w(h(e_i(\alpha)) + \epsilon) \right] \right) \cdot \tau \alpha w h'(e_i(\alpha)) \right] \quad (9)$$

Define

$$U(i) := u(V_i - e_i(\alpha)) + E v \left(\tau \cdot \left[(1 - \alpha) w/I \sum_j [h(e_j(\alpha))] + \alpha w(h(e_i(\alpha)) + \epsilon) \right] \right).$$

Consider $\sum_i U(i)$. Differentiating with respect to α gives and using (9) to cancel terms

⁶We need $E[v'(x)] + E[v''(x) \cdot x] \geq 0$ which is weaker than what we assume in the proposition.

gives

$$\begin{aligned} & \sum_i E[v'(i)] \cdot \tau \cdot (1 - \alpha) \cdot \frac{w}{I} \sum_j h'(e_j(\alpha)) \cdot e_j'(\alpha) \\ & + \sum_i E \left[v'(i) \cdot \left(-\tau \cdot \frac{w}{I} \cdot \sum_j h(e_j(\alpha)) + \tau \cdot w \cdot (h(e_i(\alpha)) + \epsilon) \right) \right]. \end{aligned}$$

We have that $E[v'(i) \cdot \epsilon] = E[v'(i)] \cdot E[\epsilon] + Cov(v'(i), \epsilon) = Cov(v'(i), \epsilon) < 0$, since $v'' < 0$. Further $\sum_i E \left[v'(i) \cdot \left(-\tau \cdot \frac{w}{I} \cdot \sum_j h(e_j(\alpha)) + \tau \cdot w \cdot (h(e_i(\alpha))) \right) \right] = \sum_i E[v'(i)] \cdot \left(-\tau \cdot \frac{w}{I} \cdot \sum_j h(e_j(\alpha)) + \tau \cdot w \cdot (h(e_i(\alpha))) \right) < 0$, since $\sum_i \left[h(e_i(\alpha)) - \frac{1}{I} \cdot \sum_j h(e_j(\alpha)) \right] = 0$ and $v'' < 0$. Further, for $\alpha = 1$, the first expression is zero. Thus for $\alpha = 1$, the derivative of the social welfare is negative, which implies that the optimal α is less than 1. $\alpha = 0$ can also not be optimal since it induces $e_i = 0$ for all i and with $h'(0) = \infty$ this violates optimality.

Example

To gain some better understanding of the assumptions and their implications we illustrate our results with an example. Let $u(x) = v(x) = -\exp(-rx)$. The FOC becomes:

$$r \exp[-r(V - e)] = \int r \{ \exp[-r\tau((1 - \alpha)W/N + \alpha w(h(e) + \varepsilon(\omega)))] \} \alpha w h'(e) f(\omega) d\omega.$$

This is equivalent to the following expression:

$$\exp[-r(V - e)] = r \{ \exp[-r\tau((1 - \alpha)W/N + \alpha w(h(e)))] \} \alpha w h'(e) \tau \int \exp[-r\tau\alpha w \varepsilon(\omega)] f(\omega) d\omega.$$

Note that the integral is the insurance effect discussed above. Now we further assume that $\varepsilon(\omega) \sim N(0, \sigma^2)$. This implies $r\tau\alpha w \varepsilon(\omega) \sim N(0, (r\tau\alpha w \sigma)^2)$. Furthermore assume $h(e) = \beta e$ for some $\beta > 0$. Using this and solving the integral, the FOC becomes:

$$\exp[-r(V - e)] = r \{ \exp[-r\tau((1 - \alpha)W/N + \alpha w(h(e)))] \} \alpha w \beta \tau \exp[(r\tau\alpha w \sigma)^2 / 2]$$

which can be simplified using $k := \ln(r\alpha w \beta \tau)$ and $R := (r\tau\alpha w \sigma)^2 / 2$ and taking logs:

$$-r(V - e) = k - r\tau[(1 - \alpha)W/N + \alpha w \beta e] + R.$$

Using $W = w(\sum e_i)\beta$ and taking care of the dropped indices yields:

$$r e_i (1 + \tau w \alpha \beta) = k + R + r \left[V_i - \tau w \beta (1 - \alpha) \frac{\sum e_i}{N} \right].$$

Using $c := r\tau w \beta (1 - \alpha)/N$ and $d := r(1 + \tau w \alpha \beta)$ gives:

$$e_i d = k + R + r V_i - c \sum_{j \neq i} e_j.$$

In matrix notation:

$$\begin{pmatrix} d & c & & c \\ c & & & \\ & & & c \\ c & & c & d \end{pmatrix} \begin{pmatrix} e_1 \\ \\ \\ e_N \end{pmatrix} = \begin{pmatrix} k + R + rV_1 \\ \\ \\ k + R + rV_N \end{pmatrix}$$

The inverse of the matrix has $\frac{d+(N-2)c}{d^2+(N-2)cd-(N-1)c^2}$ on the diagonal and $\frac{-c}{d^2+(N-2)cd-(N-1)c^2}$ off the diagonal. Note that $d > 0, c > 0, d - c > 0$. Thus the denominator is positive. We can then solve for e_i :

$$e_i = (k + R) \frac{d - c}{d^2 + (N - 2)cd - (N - 1)c^2} + r \frac{d + (N - 2)c}{d^2 + (N - 2)cd - (N - 1)c^2} V_i - r \frac{c}{d^2 + (N - 2)cd - (N - 1)c^2} \sum_{j \neq i} V_j.$$

This solution has the simple structure:

$$e_i = K_0(\alpha) + K_1(\alpha)V_i - K_2(\alpha) \sum_{j \neq i} V_j$$

Having explicitly calculated the solution for individual investments, we proceed as in the general case. We want to know how investment reacts if α rises from α^1 to $\alpha^2 > \alpha^1$. In particular, we want to replicate our cutoff result. Suppose investment falls for some j . We will check now whether it also falls for all $k > j$. Assume $\Delta e_j := e_j^2 - e_j^1 = \Delta_0 + \Delta_1 V_j - \Delta_2 \sum_{i \neq j} V_i < 0$, where $\Delta_s := K_s(\alpha^2) - K_s(\alpha^1)$. We want to show that for $k > j$ $\Delta e_k := e_k^2 - e_k^1 = \Delta_0 + \Delta_1 V_k - \Delta_2 \sum_{i \neq k} V_i < 0$. Equivalently, we want to derive conditions

under which $\frac{\partial^2 e^*}{\partial V \partial \alpha} < 0$. Note that $W/w\beta = \sum e_i = \frac{N(k+R)}{den}(d-c) + \frac{r[d+(N-2)c]}{den} - (N-1)\frac{rc}{den} = \frac{d-c}{den}[N(k+R) + r \sum V_i] > 0$.

We have $\Delta e_k - \Delta e_j = \Delta_1 \cdot (V_k - V_j) - \Delta_2 \cdot (\sum_{i \neq k} V_i - \sum_{i \neq j} V_i) = \Delta_1 \cdot (V_k - V_j) - \Delta_2 \cdot (V_j - V_k)$. Since $V_j - V_k < 0$, this term will be negative if the coefficients Δ_1 and Δ_2 are both negative. We can rewrite $\Delta_i = \int \frac{\partial \Delta_i}{\partial \alpha} d\alpha$ where the integration is from α^1 to α^2 . Tedious calculations reveal that indeed $\frac{\partial \Delta_1}{\partial \alpha} = -\frac{(N-1)\tau w \beta}{N(\tau \alpha \beta w + 1)^2} < 0$ and $\frac{\partial \Delta_2}{\partial \alpha} = -\frac{\tau w \beta}{N(\tau w \beta \alpha + 1)^2} < 0$. This proves $\Delta e_k - \Delta e_j < 0$ for $k > j$. This implies that if $\Delta e_j < 0$, then also $\Delta e_k < 0$, which is our cutoff result from the general case. Furthermore, this result implies that if $\Delta e_j > 0$ then either $\Delta e_j > \Delta e_k > 0$ or $\Delta e_j > 0 > \Delta e_k$. This result has the interpretation that richer households react less sensitively to higher investment incentives than poorer households: if the poor invest more, then the rich either invest more or less. However, if they invest more, then this is smaller than the increase in the investment of the poor. This result could not be shown in the general case, but is a nice property of our example.

Now we show that our sufficient condition about the cross partial derivative of RHS does not generally hold in our example. Calculate RHS and the relevant derivatives with respect to α and e and also the cross partial derivative:

$$RHS(e, \alpha) = r\alpha w \beta \tau \exp[-r\tau \frac{W}{N}] \exp \left\{ -r\tau \alpha \left[-\frac{W}{N} + w\beta e - \frac{1}{2} r\tau \alpha w^2 \sigma^2 \right] \right\}$$

$$\frac{\partial RHS}{\partial e} = -(r\tau\alpha w\beta)^2 \exp[-r\tau \frac{W}{N}] \exp \left\{ -r\tau\alpha \left[-\frac{W}{N} + w\beta e - \frac{1}{2}r\tau\alpha w^2\sigma^2 \right] \right\} < 0$$

$$\frac{\partial^2 RHS}{\partial \alpha \partial e} = -\alpha(rw\beta\tau)^2 \exp[-r\tau \frac{W}{N}] \exp \{ \} \left\{ 2 - \alpha r\tau \left[-\frac{W}{N} + w\beta e \right] + \frac{3}{2}(r\tau\alpha w\sigma)^2 \right\}$$

The sign of the last term is ambiguous. In the main text we assumed this mixed derivative to be negative. This holds in the example if the variance is sufficiently high so that the insurance effect dominates the redistributive effect. Recall, however, that it was shown above that the desired cutoff result holds in our example independent of the sign of this derivative.

Finally let us decompose the effect of a higher incentive parameter α on the expected marginal revenue of investment into the effects we found in the general case.

$$\frac{\partial RHS}{\partial \alpha} = rw\beta\tau \exp[-r\tau \frac{W}{N}] \exp \{ \} \left\{ 1 - r\tau\alpha \left[-\frac{W}{N} + w\beta e \right] + \frac{3}{2}(r\tau\alpha w\sigma)^2 \right\}$$

The term in the first line (which also shows up as a 1 in the brackets of the last line) is the positive incentive effect. The second term consists of the redistributive effect $\frac{W}{N} - w\beta e$ and the positive insurance effect $\frac{3}{2}(r\tau\alpha w\sigma)^2$. The general equilibrium effect was neglected. Note that the insurance effect is parameterized in the variance σ^2 .

Dynamic Problem

We show the existence and give a sufficient condition for the uniqueness of an invariant distribution of the transition probabilities of the vector e_t in the dynamic setup. This sketch shows how our analysis can be extended to a truly dynamic general equilibrium framework. This is, however, far beyond the scope of this paper.

Existence. Assume that the Jacobian matrix of (2) with respect to h is nonsingular. Now apply the implicit function theorem as stated in Hildenbrand (1974) to the equation system of FOCs (2) if $e_{t-1} \gg 0$ after writing V_i as a continuous function of ω_t, e_{t-1} with $V_i \left[(h(e_{t-1,j}, \omega_{tj}))_{j=1}^N \right] = w \cdot h(e_{t-1,i}, \omega_{t,i}) \cdot (1 - \tau)$ and endowing Ω with the metric for the discrete topology. For given ω_t, e_{t-1} this gives a unique continuous function $g_{e_{t-1}, \omega_t}(\cdot, \cdot)$ with $e_t = g_{e_{t-1}, \omega_t}(e_{t-1}, \omega_t)$ in a neighborhood of ω_t, e_{t-1} in the product topology on $[0, \bar{h}]^N$ and the discrete metric space Ω . By putting these local functions together, we find a stochastic difference equation $e_t = g(e_{t-1}, \omega_t)$, continuous in e_{t-1} , defined for $e_{t-1} \gg 0$ which determines the human capital investment in period t given the human capital investment in period $t - 1$ and the shock realization in period t . For $e_{t-1} = 0$, we know that $e_t = 0$ for all ω_t . Since as $e_{t-1} \rightarrow 0$ in some components also $g(e_{t-1}, \omega_t) \rightarrow 0$ in the same components, the function g can be continuously extended to the boundary. Then by exercise 8.10 in Stokey and Lucas (1989) [SL hereafter], the transition P generated by g according to Theorem 8.9 in SL has the Feller property. Thus, since the state space $S = [0, \bar{h}]^N$ is compact, by Theorem 12.10 in SL, an invariant distribution of P exists.

Uniqueness. To apply Theorem 12.12 from SL, the mixing condition given in assumption 12.1 in SL and the monotonicity of P have to be checked.

References

- [1] Abel, Andrew B. (1999), The Social Security Trust Fund, the Riskless Interest Rate, and Capital Accumulation, NBER Working Paper No. W6991.
- [2] Bohn, Henning (1998), Risk Sharing in a Stochastic Overlapping Generations Economy, University of Santa Barbara Discussion Paper, December 1998).
- [3] Bohn, Henning (2000), Social Security and Demographic Uncertainty: The Risk Sharing Properties of Alternative Policies, in: John Y. Campbell and Martin Feldstein (eds.), Risk Aspects of Investment-Based Social Security Reform, Chicago: The University of Chicago Press, 2000.
- [4] Boldrin, Michele and Ana Montes (2002), The Intergenerational State. Education and Pensions, mimeo.
- [5] Cigno, Alessandro, Annalisa Luporini and Anna Pettini (2003), Transfers to families with children as a principal-agent problem, *Journal of Public Economics* 87, 1165-1177.
- [6] Demange, Gabrielle and Guy Laroque (1999), Social Security and Demographic Shocks, *Econometrica* 67, 527-42.
- [7] Enders, Walter and Harvey E. Lapan (1982), Social Security Taxation and Intergenerational Risk Sharing, *International Economic Review* 23, 647-658.
- [8] Gollier, Christian (2001), *The Economics of Risk and Time*, MIT Press.
- [9] Gordon, Roger H. and Hal R. Varian (1988), Intergenerational Risk Sharing, *Journal of Public Economics* 37, 185-202.
- [10] Heckman, James J. (1999), Policies to Foster Human Capital, University of Chicago, Working Paper.
- [11] Hildenbrand, Werner (1974), *Core and Equilibria of a Large Economy*, Princeton: Princeton University Press.
- [12] Holmstrom, Bengt (1979), Moral Hazard and Observability, *Bell Journal of Economics* 10, 74-91.
- [13] Holmstrom, Bengt (1982), Moral Hazard in Teams, *Bell Journal of Economics* 13, 324-240.
- [14] Kolmar, Martin (1997), Intergenerational Redistribution in a Small Open Economy with Endogenous Fertility, *Journal of Population Economics* 10, 335-56.
- [15] Merton, Robert S. (1983), On the Role of Social Security as a Means for Efficient Risk Sharing in an Economy Where Human Capital Is Not Tradable, in: Zvi Bodie and John B. Shoven (eds.), *Financial Aspects of the United States Pension System*, 325-358.
- [16] Mirrlees, James (1974), Notes on Welfare Economics, Information and Uncertainty, in: M. Balch, D. McFadden and S. Wu (eds.), *Essays in Economic Behavior under Uncertainty*, 243-258.
- [17] Mookherjee, Dilip (1984), Optimal Incentive Schemes with Many Agents, *Review of Economic Studies* 51, 433-46.
- [18] Schweizer, Urs (1996), Endogenous Fertility and the Henry George Theorem, *Journal of Public Economics* 61, 209-28.

- [19] Shiller, Robert (1999), Social Security and Intergenerational, Intragenerational and International Risk Sharing, *Carnegie-Rochester Series in Public Policy* 50, 165-204.
- [20] Smith, Alasdair (1982), Intergenerational Transfers as Social Insurance, *Journal of Public Economics* 19, 97-106.
- [21] Stokey, Nancy and Robert E. Lucas (1989), *Recursive Methods in Economic Dynamics*, Cambridge, Mass. and London: Harvard University Press.
- [22] Topkis, Donald M. (1998), *Supermodularity and Complementarity*, Princeton: Princeton University Press.

Correspondence:
Ashok Kaul
Gutenberg University Mainz
Welderweg 4
D-55099 Mainz
Germany
email: ashok.kaul@uni-mainz.de