INVESTMENT IN CHILDREN, SOCIAL SECURITY, AND INTRAGENERATIONAL RISK SHARING

Simon Fan, Yu Pang, Pierre Pestieau

REPRINT | 3163

UCLouvain

CORE | LIDAM
Louvain Institute of Data Analysis and Modeling in economics and statistics

CORE

Voie du Roman Pays 34, L1.03.01

B-1348 Louvain-la-Neuve

Tel (32 10) 47 43 04

Email:lidam-library@uclouvain.be

https://uclouvain.be/en/research-institutes/lidam/core/core-reprints.html

International Tax and Public Finance
https://doi.org/10.1007/s10797-021-09664-3

Investment in children, social security, and intragenerational risk sharing

Simon Fan[1] · Yu Pang[2] · Pierre Pestieau[3]

Accepted: 13 February 2021
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2021

Abstract
We analyze the role of pay-as-you-go social security in intragenerational risk sharing in an overlapping-generations model with individual heterogeneity. Parents invest in their children's education in state schools in exchange for old-age financial support. Due to random factors such as luck in the job market, children may have different earning capacities despite that they receive the same education. Without social security, a parent gets a transfer payment from her own child, so the received amount is uncertain as it depends on the child's earnings. The social security scheme, which essentially serves to pool transfer contributions from all children and then redistribute them equally to each parent, insures parents against the risk of educational investments. Our model shows that social security stimulates educational spending, enhances labor earnings, and increases *ex ante* individual utility. However, it may worsen *ex post* intragenerational inequality of lifetime income.

Keywords Old-age insurance · Educational investment · Social security · Inter-family risk pooling · Income inequality

JEL Classification D81 · H20 · H55 · I24

Extend your filial piety for your aged parents to all the aged, and extend your care for your own children to all children.
— Mencius (372 BC – 289 BC, the "Second Sage" of Confucianism)

✉ Pierre Pestieau
P.Pestieau@ulg.ac.be

Simon Fan
fansimon@ln.edu.hk

Yu Pang
ypang@must.edu.mo

1 Department of Economics, Lingnan University, Tuen Mun, Hong Kong

2 School of Business, Macau University of Science and Technology, Taipa, Macau

3 Department of Economics, University of Liège and CORE, Liège, Belgium

Published online: 29 March 2021

1 Introduction

People have reared and educated children in exchange for their old-age support for most of human history. Cheung (1972, p. 641) describes the parent–child relations in ancient China as "[j]ust as dogs were raised to hunt for their masters before they were pets, so in early traditional China children were raised as a source of income and a store of wealth." This is still the case in many developing countries. Such an implicit intrafamily contract can be sustained by intergenerational trade (Ehrlich and Lui 1991, 1998) and by social norms (Becker 1993).[1] In developed economies, however, the intrafamily contract has generally been displaced by the Pay-As-You-Go (PAYGO) social security program, which calls for the young to finance retirement benefits for the old.

A core function of social security is to manage the risks of random events leaving the elderly impoverished (Diamond 1977; Shiller 1999).[2] The high transaction costs associated with the operation of private insurance markets means that there is a legitimate role for government-administered social protection (Fleurbaey 2008; Cremer and Roeder 2017; Pestieau and Lefebvre 2018). Most theoretical studies on the risk management function of social security focus on the intertemporal resolution of financial uncertainty.[3] They suggest that PAYGO is an effective means of pooling aggregate risks across generations: the old obtain financial aid from the young if they suffer a setback and save the young from the burden of pensions if they experience a favorable shock (Krueger and Kubler 2002; Shiller 2003).

This paper analyzes *intra*-generational risk sharing in the presence of *inter*-generational links. Working-age population may have different earning capacities even if they receive the same education at school. Their wage differences are shaped by some non-educational random factors such as health, personality, and fortune (e.g., Brown 1980; Bowles et al. 2001). Hartog et al. (2007) estimate that *ex ante* risk of university education is comparable to that of a randomly selected financial portfolio of 30 stocks listed on the New York Stock Exchange. If different generations are linked through wealth transfers, the random factors that matter for every member of workforce tend to shape not only her own consumption/investment decision but also the size of her ascending transfer contribution. The existing literature pays relatively little attention to the impact of social security on intragenerational risk sharing. Notable exceptions include Merton (1983), Storesletten et al. (1999) and Nishiyama

[1] Ehrlich and Lui (1991) establish a model in which parents invest in their children's education and expect financial support from their children when they become old. This parent–child interaction is vividly described as a mutually productive "intergenerational trade."

[2] Social security also serves some other important purposes, such as redistributing wealth (Boadway and Marchand 1995; Deaton et al. 2002), giving employees more leisure time (Cremer et al. 2007), promoting growth in poor countries (Boadway 2006; Glomm and Kaganovich 2008), encouraging life-cycle savings (Bloom et al. 2007), reducing family formation and fertility (Ehrlich and Kim 2007), tackling the externality associated with fertility and human capital accumulation (Cremer et al. 2011), and attenuating the response of career lengths to earnings shocks (Ljungqvist and Sargent 2014).

[3] See Enders and Lapan (1982), Gordon and Varian (1988), Fuster et al. (2003) and Ball and Mankiw (2007).

and Smetters (2007), but they abstract from the consideration of the intergenerational link via wealth transfers.

We establish an overlapping-generations (OLG) model in which individuals receive education in childhood, earn a wage and bear one child in young adulthood, and retire and obtain an intergenerational transfer in old adulthood. We assume that individuals are *ex ante* identical and receive public education only. In the labor market, lucky and unlucky individuals become *ex post* high and low earners, respectively. To afford their old-age consumption, young adults have two investment options: (1) savings (e.g., bank deposit), which earn a constant but modest interest, and (2) funding public education to invest in children's human capital, which raises children's future earnings (and thus transfer contributions) but exhibits diminishing returns.

In Sect. 3, we examine the transitional dynamics and steady state of the optimal investment decisions under two systems of intergenerational transfers. In the first, each young individual gives a fixed fraction of her earnings to her old parent in compliance with a social norm.[4] As the transfer occurs within a family, the returns to educational investments are volatile: a parent gets a sizeable transfer if her child is offered a high-paying job in the future but little otherwise. This perceived risk is a major concern in people's decision making processes, which may cause them to prefer riskless savings over risky educational investments.

Yet the concern will be relieved after the introduction of the second type of intergenerational transfer—the PAYGO social security scheme, which enables the old generation to obtain the same pension funded by the young generation. Social security insures people against the risk tied to their children's future earnings; those whose children have a limited earning capacity receive old-age support from the high-paid children of other families. By pooling the risks of retirement benefits among the whole next generation, social security eliminates the *ex post* variability of returns to the educational investment, making it as risk-free as a bank deposit. Consequently, more investments will be made in children's schooling, which fuels human capital accumulation of the economy.

Section 4 shows that social security creates two welfare gains compared with the intrafamily contract. First, under the social security scheme, individuals become insured against fluctuations in the received intergenerational transfer, which improves their ability to smooth their consumption. Second, social security encourages human capital development and hence increases labor earnings. Our analysis provides a new rationale as to why social protection often accounts for a large share of government expenditures in developed countries (OECD 2020) and why Nordic countries adopt the system of welfare state.

The next question arises as to how the introduction of social security affects income inequality? We present a surprising result that under the public education system, social

[4] A social norm that stipulates people to transfer a fixed fraction of their wages to support their parents' old-age consumption can be sustained in a Nash equilibrium. Anyone who violates the social norm will expect that his child will leave him uncared in his old age, deterring deviations. See Ehrlich and Lui (1991) and Becker (1993).

security increases the inequality of a generation's lifetime disposable incomes. Without social security, individuals may move up or down the income ladder in old age. A low-paid worker will attain a generous transfer if her child becomes a high earner in future, but a high-paid worker will get little if his child has a low earning capacity. In contrast, redistributive social security equalizes everyone's pension, hindering intragenerational income mobility. A polarity in income thus persists as high-paid workers remain richer than their low-paid peers over the life cycle.

In Sect. 5, we extend our model to characterize the *optimal* intergenerational contract, namely the fraction of children's earnings that should be transferred to parents to maximize the steady-state social welfare. We obtain an analytical solution to the optimal contract under the social security scheme and find that it is unaffected by intertemporal consumption preferences but decreases with the educational efficiency in certain circumstances. We then compare the optimal intergenerational contract with and without social security in a quantitative simulation, which demonstrates that the introduction of social security leads to a large size of ascending transfer.

Despite the focus on social security, our paper also complements the literature on inter-family risk sharing. Lucas (1992) builds a model characterizing the optimal allocation of resources when infinitely-lived households experience stochastic, idiosyncratic preference shocks that affect their urgency to consume. Townsend (1994) finds that in poverty-stricken Indian villages, gifts in family networks and credit transactions smooth adverse fluctuations in incomes caused by diseases and erratic weather. Fafchamps and Lund (2003) show that villagers receive gifts and informal loans from their friends and relatives to smooth income and expenditure shocks in rural Philippines. Our paper addresses a new angle of inter-family risk sharing via the pooling of uncertainty associated with children's future earnings and transfer payments.

2 The model setup

Consider an overlapping-generations economy populated by a unit mass of individuals who live for three periods: childhood, young adulthood, and old adulthood. Children receive education and do not make any decisions. Young adult bear one child, earn a wage, choose their own consumption and saving levels, and contribute part of wages to their parents. Old adults retire and live on savings and transfer payments without leaving bequests. Individuals who spend young adulthood in period t are labeled as "generation t."

In period $t-1$, the government provides every child (i.e., generation t) with public education, of which the expenditure, e_{t-1}, is equally shared among all the parents (i.e., generation $t-1$). As children grow up in period t, their human capital accrues to $H(e_{t-1})$, which satisfies

$$H' > 0, \quad H'' < 0, \quad \lim_{e \to 0} H'(e) = \infty. \tag{1}$$

The first two conditions indicate that educational investments exhibit positive but diminishing returns, while the last condition ensures that $e > 0$ in equilibrium.

Following Glomm and Ravikumar (1992) and de la Croix and Doepke (2004, 2009), we assume that individuals who receive the same public education are *ex ante* identical in childhood. However, they earn different wages after entering the job market. To determine the wage of a representative member of generation t (w_t), we follow Pestieau and Possen (1979), who consider that the wage depends on her human capital, $H(e_{t-1})$, and the job market uncertainty, ε:

$$w_t = (1 + \varepsilon)H(e_{t-1}), \tag{2}$$

where ε is a non-educational determinant of one's earning capacity such as fortune.[5] For simplicity, assume that ε is an independent and identically distributed random component, taking either of the following two values in each period:

$$\varepsilon = \begin{cases} \varepsilon_h > 0 & \text{with a probability of } p \\ \varepsilon_l \in (-1, 0) & \text{with a probability of } 1 - p \end{cases}. \tag{3}$$

Equation (3) means that an individual has a good luck on the job market and becomes rich (h-type) if $\varepsilon = \varepsilon_h$, but she has an ill luck and becomes poor (l-type) if $\varepsilon = \varepsilon_l$. By the law of large numbers, each generation comprises a fraction p of h-type members and a fraction $1 - p$ of l-type members. Assume that job perspective is uncertainty neutral so that the expected value of ε is zero:

$$E(\varepsilon) = p\varepsilon_h + (1 - p)\varepsilon_l = 0. \tag{4}$$

In period t, the interactions between the government and members of generation t proceed in two stages. The government moves first to choose the (public) educational spending per child, e_t, to maximize the social welfare, which is defined as individual expected utility under utilitarianism:

$$U_t = pU_{ht} + (1 - p)U_{lt}, \tag{5}$$

where U_{it} for $i \in \{h, l\}$ is the lifetime utility of an i-type member of generation t. After observing the government's decision, each i-type member chooses maximizes her utility as specified by

$$U_{it} = u(c_{it}) + v(d_{it}), \tag{6}$$

where c_{it} and d_{it} denote her young-age and old-age consumption levels, and utility functions are increasing and concave (i.e., $u' > 0, v' > 0, u'' < 0, v'' < 0$). The budget constraints imply that

[5] For example, Arteaga (2018) finds evidence that an employee's wage is positively related to her human capital and her employer's impression of her performance during the recruitment process. Pluchino et al. (2019) create a computer model to predict that luck is more important than talent for determining one's career success.

$$c_{it} = (1-\delta)[(1+\varepsilon_i)H(e_{t-1})] - s_{it} - e_t \quad d_{it} = (1+r)s_{it} + f_t, \tag{7}$$

where s_{it} is the saving of an i-type member of generation t, δ is the fraction of her earnings transferred to the previous generation, r is the (exogenous) interest rate, and f_t is the intergenerational transfer (retirement benefit) received by a member of generation t.

We compare two schemes of transferring wealth to retirees. First, absent social security, each member of generation t receives financial support in period $t+1$ from her own child, who earns a wage of $(1+\varepsilon)H(e_t)$ where ε denotes the child's type. In other words, every young adult allocates a fraction δ of her wage to support her parent's old-age consumption, and she expects her child to do the same when she becomes old. Second, under the PAYGO pension scheme, every young adult pays a flat social security tax amounting to a fraction δ of her wage to the pension funds, which are then equally shared by the old generation.[6] In sum, f_t is written as

$$f_t = \begin{cases} \delta(1+\varepsilon)H(e_t) & \text{if there is no social security scheme} \\ \delta E[(1+\varepsilon)H(e_t)] = \delta H(e_t) & \text{if there is a social security scheme} \end{cases}, \tag{8}$$

which increases with e_t. In other words, people view children's schooling as an investment vehicle: the more they spend in educating children, the higher returns they gain at the retirement age.

Before deriving the equilibrium outcome of the model, we develop the next lemma to compare the savings between h-type and l-type members of the same generation:

Lemma 1 *An h-type individual saves more than an l-type individual* ($s_{ht} > s_{lt}$ all for t).

Proof See Appendix 1. □

Lemma 1 suggests the rich's saving level is higher than the poor's. This result is supported by many empirical findings. Dynan et al. (2004) demonstrate a significantly positive relationship between saving and income using data on American households. Bozio et al. (2017) and Gandelman (2017) show similar evidence in their respective studies of England and Latin America.

[6] In practice, a retiree's pension may depend on a range of factors such as her age, marital status, and tax payment. For example, the poor receive substantial retirement benefits even though they contribute little in young age, while the rich get back only a small fraction of tax contributions through pensions. Our stylized model addresses the redistributive mechanism of social security, as it works in many developed countries. Given the key role of social security in wealth redistribution, our results will hold qualitatively even when more practical issues are taken into account.

3 Equilibrium dynamics: no social security versus social security

In this section, we characterize the transitional dynamics and steady state of the economy. We first consider the economy where social security is absent and intergenerational transfers are made in accordance with an intrafamily contract. We then explore how the outcome changes after social security is introduced. To provide a comparison benchmark, we assume that children give a fixed fraction of their earnings to their parents regardless of the presence of social security (namely, δ is a parameter) in this section and the next. This assumption will be relaxed in Sect. 5 where the optimal transfers under the two schemes are examined.

3.1 No social security scheme

Consider the case where an individual receives old-age support from her own child. In young adulthood, each individual has two investment options. She can either deposit the money into a bank or invest in public education to increase her child's human capital. While the interest paid on a deposit is a constant r, educational investments exhibit diminishing returns (i.e., $H'' < 0$). More importantly, compared with risk-free savings, educational investments are risky in that the return depends critically on the child's luck in the job market.

The interactions between the government and young adults (as described in Section 2) can be analyzed by backward induction. An i-type member of generation t expects her child to become h-type with a probability p and l-type with a probability $1-p$ in period $t+1$. Given the educational expenditure in period t, she chooses her saving level to maximize her utility as expressed by

$$\begin{aligned} U_{it}^n = {} & u[(1-\delta)(1+\varepsilon_i)H(e_{t-1}^n) - s_{it}^n - e_t^n] \\ & + pv[(1+r)s_{it}^n + \delta(1+\varepsilon_h)H(e_t^n)] \\ & + (1-p)v[(1+r)s_{it}^n + \delta(1+\varepsilon_l)H(e_t^n)], \end{aligned} \tag{9}$$

where the superscript n stands for *no social security*. The first-order condition of (9) obtains

$$\begin{aligned} & (1+r)\big\{pv'[(1+r)s_{it}^n + \delta(1+\varepsilon_h)H(e_t^n)] \\ & \quad + (1-p)v'[(1+r)s_{it}^n + \delta(1+\varepsilon_l)H(e_t^n)]\big\} \\ & = u'[(1-\delta)(1+\varepsilon_i)H(e_{t-1}^n) - s_{it}^n - e_t^n], \end{aligned} \tag{10}$$

where s_{it}^n is a function of (e_{t-1}^n, e_t^n).

Taking into account the composition of the working-age population and each i-type worker's saving, the government chooses e_t^n to maximize the social welfare in (5), which can be rewritten as

$$
\begin{aligned}
U_t^n = p\big\{&u[(1-\delta)(1+\varepsilon_h)H(e_{t-1}^n) - s_{ht}^n - e_t^n] \\
&+ (1-p)v[(1+r)s_{ht}^n + \delta(1+\varepsilon_l)H(e_t^n)] \\
&+ pv[(1+r)s_{ht}^n + \delta(1+\varepsilon_h)H(e_t^n)]\big\} \\
&+ (1-p)\big\{u[(1-\delta)(1+\varepsilon_l)H(e_{t-1}^n) - s_{lt}^n - e_t^n] \\
&+ (1-p)v[(1+r)s_{lt}^n + \delta(1+\varepsilon_l)H(e_t^n)] \\
&+ pv[(1+r)s_{lt}^n + \delta(1+\varepsilon_h)H(e_t^n)]\big\}.
\end{aligned} \tag{11}
$$

The government's optimal educational expenditure per child in period t satisfies

$$
\frac{\mathrm{d}U_t^n}{\mathrm{d}e_t^n} = 0. \tag{12}
$$

Based on (12), we establish the following proposition featuring the transitional dynamics of educational expenditure per child in absence of social security:

Proposition 1 *Without the social security scheme, the educational expenditure per child changes over time, and satisfies*

$$
\forall\, t, \quad \frac{1+r}{\delta} < H'(e_t^n) < \frac{1+r}{\delta(1+\varepsilon_l)}. \tag{13}
$$

Proof See Appendix 1. □

Proposition 1 shows that e_t^n is time dependent, i.e., e_t^n is related to e_{t-1}^n. Besides, the solution to e_t^n falls in the specific range as shown in (13). With e_t^n being solved in (12), we can proceed to solve the model. For illustrative purpose, we provide a numerical example with specific functional forms of $u(.)$, $v(.)$, and $H(.)$ in Appendix 2 illustrating the evolutions of $(e_t^n, s_{lt}^n, s_{ht}^n, U_t^n)$ for $t = 0, 1, 2, \ldots$. In the long run, e_t^n will converge toward a constant level, and so will the other three variables. We define the steady state of the economy as that the education spending per child remains fixed over time:

Definition 1 The steady state of the economy is determined by the time-invariant educational expenditure per child (i.e., $e_t = e$ for all t).

Based on Definition 1, we can pin down the steady-state public educational spending per child and private saving levels by Eqs. (10) and (12). It follows that the average labor income amounts to $H(e)$ in the steady state (see Eq. (2)).

3.2 Social security scheme

We now investigate the role of the PAYGO social security scheme in determining educational spending and private savings. Under the social security scheme, each old adult receives the average transfer contributed by the young adults rather than

her own child's transfer. In other words, social security allows the risk of educational investments to be spread over the entire next generation.

To examine the interactions between the government and generation t, we insert (7)–(8) into (6) yielding the utility of an i-type member of generation t:

$$U_{it}^k = u\left[(1-\delta)(1+\varepsilon_i)H(e_{t-1}^k) - s_{it}^k - e_t^k\right] + v\left[(1+r)s_{it}^k + \delta H(e_t^k)\right], \tag{14}$$

where we use the superscript k to denote *social security*. An i-type individual chooses her saving, s_{it}^k, to maximize U_{it}^k according to the following first-order condition:

$$u'\left[(1-\delta)(1+\varepsilon_i)H(e_{t-1}^k) - s_{it}^k - e_t^k\right] = (1+r)v'\left[(1+r)s_{it}^k + \delta H(e_t^k)\right]. \tag{15}$$

Substituting (14) into (6) yields the social welfare in period t as

$$\begin{aligned} U_t^k = {} & p\left\{u[(1-\delta)(1+\varepsilon_h)H(e_{t-1}^k) - s_{ht}^k - e_t^k] + v[(1+r)s_{ht}^k + \delta H(e_t^k)]\right\} \\ & + (1-p) \\ & \left\{u[(1-\delta)(1+\varepsilon_l)H(e_{t-1}^k) - s_{ht}^k - e_t^k] + v[(1+r)s_{lt}^k + \delta H(e_t^k)]\right\}. \end{aligned} \tag{16}$$

Maximizing (16) with respect to e_t^k yields the interior solution to the government's optimal educational expenditure, as specified in the next proposition:

Proposition 2 *Under the social security scheme, the optimal educational expenditure per child is determined by*

$$\forall\, t, \quad H'(e_t^k) = \frac{1+r}{\delta}. \tag{17}$$

Proof See Appendix 1. □

As social security is introduced, funding public schools becomes as safe an investment option as depositing money at a bank. By investing e_t^k in children's education, a representative member of generation t expects the marginal return to be $\delta H'(e_t^k)$. Equation (17) implies that in equilibrium this return is equal to the gross benefit from personal savings, $1+r$. Provided that $H'' < 0$, it is straightforward to show that e_t^k decreases with r but increases with δ. Intuitively, a higher interest rate (larger r) increases the attractiveness of savings and hence the opportunity cost of educational investment. If children are required to transfer a greater fraction of their wages to their parents (a larger δ), then self-interested parents will be motivated to better educate children so that they will have a higher level of human capital and earn a higher wage in the future (i.e., $H' > 0$).

Note that e_t^k in (17) differs from e_t^n in (12) in at least three aspects. First, e_t^k is independent of the job market uncertainty (ε). Under the social security scheme, an individual's retirement benefit is funded by all children, which makes educational investments as safe as savings. Second, e_t^k does not depend on utility functions (u' and v'). The individual decision-making involves a trade-off between two investment options that gain determinate returns, but the equilibrium will not be affected by

their consumption preferences. Third, e_t^k keeps constant over time (i.e., $e_t^k = e^k$ for all t), exhibiting no transitional dynamics. Since the economy always operates at steady state, the average wage is constant at $H(e^k)$ for all generations. Each generation contributes part of its wages to the previous generation in young adulthood and is compensated in retirement with an equal amount by the next generation, which ultimately leaves their lifetime incomes unchanged.

With e_t^k being analytically solved in (17), an interior solution to s_{it}^k in (15) can be derived accordingly. In particular, given that e_t^k is unrelated with ε_i (Proposition 2), the relationship between s_{it}^k and ε_i can be obtained from the total differentiation of (15) with respect to ε_i and s_{it}^k as follows:

$$\frac{\mathrm{d}s_{it}^k}{\mathrm{d}\varepsilon_i} = \frac{(1-\delta)H(e_{t-1}^k)u''}{(1+r)^2 v'' + u''} > 0. \tag{18}$$

This positive response of s_{it}^k to ε_i captures the typical "income effect": an increased wage induces individuals to save more for a higher level of old-age consumption.

4 A comparative analysis: impacts of social security

We have so far solved for the equilibrium with and without social security. In this section, we compare the two equilibrium outcomes and demonstrate that social security has an edge in reducing the investment risk, encouraging human capital accumulation, and promoting social well-being. Our analysis complements the literature by showing that social security is conducive to economic development through its role of inter-family risk sharing. We also investigate the influence of social security on income inequality.

To measure the risk associated with generation t's returns from investing in their children, we rely on the coefficient of the variation of f_t. The following proposition discusses the riskiness of educational investments:

Proposition 3 *Educational investment incurs a risk of* $\sqrt{-\varepsilon_l \varepsilon_h}$ *without social security but incurs no risk under the social security scheme.*

Proof See Appendix 1. □

Individuals who receive old-age support from their own children expect their educational investments to earn $\delta(1+\varepsilon_h)H(e^n)$ with a probability p and $\delta(1+\varepsilon_l)H(e^n)$ with a probability $1-p$. Under the contract of intrafamily transfer, educational investments entail a risk of $\sqrt{-\varepsilon_l \varepsilon_h}$. Such a risk will be eliminated by the implementation of the PAYGO social security scheme, which guarantees the return to educational investments. Clearly, the introduction of social security tends to reward parents for building children's human capital, as demonstrated in the next proposition:

Proposition 4 *Social security raises the educational expenditure per child* $(e^k > e_t^n)$.

Proof See Appendix 1. □

Proposition 4 suggests that compared with the intrafamily contract, the PAYGO social security program improves children's educational attainment. To protect themselves against the riskiness of educational investments when social security is absent, individuals allocate more of their funds to riskless savings. However, social security removes the risks of educational investments, making it more attractive for the current generation to develop the next generation's human capital as that will increase their own retirement benefits. Our result that social security has a positive effect on education is consistent with the literature (e.g., Echevarría and Iza 2006).

In our model, education directly fosters human capital formation, which is the only production factor generating income. We hence infer from Proposition 4 that social security enhances production and raises labor earnings. Empirical research has offered supportive evidence. Bellettini and Ceroni (2000) estimate that social protection has a positive impact on human capital formation and incomes using both cross-section and panel data. Lee and Chang (2006) find that social security boosts economic development through human capital accumulation in OECD countries.

Although the implementation of social security induces educational investments, the investment efficiency may not be fully exploited. According to Baland and Robinson (2000), an efficient educational investment satisfies

$$H'(e_t^*) = 1 + r. \tag{19}$$

The next proposition follows directly from Proposition 4 and Eq. (19):

Proposition 5 *Both ways of intergenerational transfer lead to an inefficient educational investment, but the implementation of social security improves efficiency, i.e.,* $e_t^n < e_t^k < e_t^*$.

Proof See Appendix 1. □

Proposition 5 predicts that when old-age support is provided by the next generation, educational investments remain inefficient no matter whether or not social security is present. This result holds even if saving generates a negative return. Because parents obtain only part of children's incomes, they make investment decisions without accounting for the other part, which is at the children's disposal. This constraint limits the efficiency of allocating economic resources. Nevertheless, the introduction of the PAYGO social security scheme tends to enhance the efficiency of the investment in human capital.

We proceed to discuss the welfare impact of social security in the following proposition:

Proposition 6

(i) *Given the educational spending, social security improves social welfare, i.e.,* $U^k(e_t) > U^n(e_t)$;

(ii) *Social security improves social welfare in equilibrium, i.e.,* $\max\{U_t^k\} > \max\{U_t^n\}$.

Proof See Appendix 1. □

Proposition 6(i) suggests that the introduction of social security is socially desirable even when its stimulating effect on educational investments is not taken into account. The upside of social security is to eliminate the uncertainty in an individual's old-age consumption. Luttmer and Samwick (2018) find that Americans are averse to the perceived uncertainty in their retirement benefits; such uncertainty is costly in that it hinders their ability to smooth consumption. By the same logic, social security increases *ex ante* individual utility through its role as old-age insurance.

Social security's stimulating effect on educational investments reinforces its welfare-enhancing impact in equilibrium. Proposition 6(ii) predicts that the presence of social security always makes people better off in equilibrium. This result adds to the theoretical literature on the welfare analysis of social security. Sinn (2004) suggests the PAYGO pension system provides insurance against not having children and against children's ungratefulness to their old parents. Cremer et al. (2006) posit that social security can promote social well-being by insuring parents against stochastic fertility. Empirically, Drèze and Khera (2017) find that the expansion of social security programs make a great contribution to Indian people's welfare.

While social security improves *ex ante* social welfare, it also affects *ex post* income inequality. For each generation, income inequality is measured based on the distribution of members' lifelong disposable incomes, which refer to their wages minus transfer payments plus retirement benefits. To gauge the degree of economic inequality, we rely on the Gini coefficient.

Given the composition of h-type and l-type members, each generation can be divided into four cohorts – a fraction $(1-p)^2$ of l-type members with l-type children, a fraction $p(1-p)$ of l-type members with h-type children, a fraction $p(1-p)$ of h-type members with l-type children, and a fraction p^2 of h-type members with h-type children. Absent social security, an i-type individual of generation t with a j-type child, where $i, j \in (l, h)$, has a lifetime disposable income at the level of

$$I_{ijt}^n = (1-\delta)(1+\varepsilon_i)H(e_{t-1}^n) + \delta(1+\varepsilon_j)H(e_t^n). \tag{20}$$

In the steady state, the lifetime income of each cohort can be rewritten as

$$\begin{aligned} I_{ll}^n &= (1+\varepsilon_l)H(e^n), & I_{lh}^n &= [1+(1-\delta)\varepsilon_l + \delta\varepsilon_h]H(e^n), \\ I_{hl}^n &= [1+(1-\delta)\varepsilon_h + \delta\varepsilon_l]H(e^n), & I_{hh}^n &= (1+\varepsilon_h)H(e^n). \end{aligned} \tag{21}$$

Because $\varepsilon_h > \varepsilon_l$, it is straightforward to show the income ranking as $I_{ll} < \{I_{lh}, I_{hl}\} < I_{hh}$, where $I_{lh} < I_{hl}$ if and only if $\delta < \frac{1}{2}$. Cohorts with incomes I_{lh} or I_{hl} can be viewed as "middle-class".

Figure 1a provides a visual aid to construct the Gini coefficient without social security. Each of the four segments of the Lorenz curve is a straight line, meaning

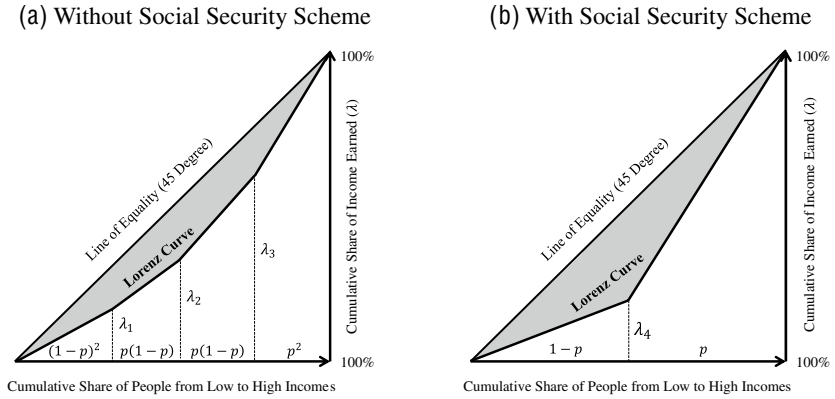

Fig. 1 The Gini coefficient with and without social security. *Note*: The Lorenz curve depicts the share of income cumulatively earned by the bottom population percentile. The Gini is the ratio of the shaded area to the area below the 45-degree line. See the expressions of λ_1–λ_4 in Appendix 1

income equality within a cohort; however, each segment has a different slope, signifying income inequality between cohorts. With λ_1–λ_3 representing the proportions of total income earned by the poorest cohort, the poorest two cohorts, and individuals other than the richest cohort, we compute the Gini coefficient as

$$\begin{aligned} G^n &= 1 - 2 \\ &\left[\frac{(1-p)^2\lambda_1}{2} + \frac{p(1-p)(\lambda_1+\lambda_2)}{2} + \frac{p(1-p)(\lambda_2+\lambda_3)}{2} + \frac{p^2(\lambda_3+1)}{2}\right] \\ &= 1 - p^2 - (1-p)\lambda_1 - 2p(1-p)\lambda_2 - p\lambda_3. \end{aligned} \tag{22}$$

With social security, a generation can be classified with respect to their lifetime income into two groups—a proportion p of h-type members with an income of $(1+\varepsilon_h)H(e^k)$ and a proportion $1-p$ of l-type members with an income of $(1+\varepsilon_l)H(e^k)$, where $\varepsilon_h > \varepsilon_l$. Figure 1b plots that under social security scheme, the Lorenz curve is composed of two straight lines. With λ_4 being the proportion of income earned by all l-type members, we calculate the Gini coefficient as

$$G^k = 1 - 2\left[\frac{(1-p)\lambda_4}{2} + \frac{p(\lambda_4+1)}{2}\right] = 1 - p - \lambda_4. \tag{23}$$

By comparing (22) and (23), we discuss the Gini coefficient in the following proposition:

Proposition 7 *In the steady state, lifetime incomes are less equal under social security scheme than under intrafamily transfer scheme, i.e.,* $G^n < G^k$.

Proof See Appendix 1. □

Proposition 7 suggests that social security may propagate income inequality. Intergenerational intra-family transfers give rise to the possibility of social mobility over the life cycle. A low-wage worker gets a large transfer payment in retirement if her child grows up to be a high earner, while a high-wage young adult receives little retirement benefit if her child works in a low-paying job. Such intragenerational mobility, however, is blocked by the social security that makes the whole generation's retirement benefits equalized. This eliminates the "middle-class," with members of each generation either rich or poor, resulting in a more uneven income distribution. Our result is consistent with Gokhale et al. (2001) calibration finding that were it not for social security, the U.S. income inequality would be lower.

One caveat needs noting: Proposition 7 follows largely from the assumption of equal (public) education for every child. The literature often addresses that a person's human capital is governed by her parental human capital (instead of income) and received educational spending (de la Croix and Michel 2002). For example, Shea (2000) finds that for most families a shock in parental income has little impact on child's human capital. In our highly stylized model, children receive the same education, and also they have the same parental human capital because their parents receive the same education. Consequently, they have the same level of human capital and the expected income. If we instead consider that the rich spend more than the poor on children's education (which is often the case), then we need to be more cautious about the prediction of Proposition 7.

5 The optimal intergenerational contract

In the previous two sections, we have analyzed the equilibrium dynamics taking the proportion of transfer as given (i.e., exogenous δ) and found a key role of δ in shaping the equilibrium outcome and social welfare. This gives rise to the question of what the optimal size of transfer is. With too generous retirement benefits (large δ), young adults live on a tight consumption budget. They are not only overexploited to provide financial support for their parents but also incentivized by the high return on educational investments to devote more to their children's schooling. With overly low retirement benefits (small δ), the expectation of obtaining little old-age support from children discourages educational spending and thus undermines economic development.

In this section, we characterize from the social planner perspective the intergenerational transfer that maximizes the steady-state social welfare, namely, $\widetilde{\delta} := \arg\max\{U\}$. For mathematical tractability of the maximization problem, we consider the following specific functions:

$$u(c_t) = \ln c_t, \quad v(d_t) = \alpha \ln d_t, \quad H(e_t) = \beta \ln e_t, \tag{24}$$

where $\alpha > 0$ indicates the preference for old-age consumption, and $\beta > 0$ measures the return on children's education.

With social security, we write the steady-state welfare in (16) as $U^k(\delta)$. The next proposition characterizes the welfare-maximizing intergenerational contract, $\widetilde{\delta}^k$:

Proposition 8 *$\widetilde{\delta}^k$ is determined by*

$$\frac{(r+\varepsilon_h+r\varepsilon_h)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta}{(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta} = \frac{\varepsilon_h}{\varepsilon_l}\cdot\frac{[1+(r+\varepsilon_h+r\varepsilon_h)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta}{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta}. \tag{25}$$

Proof See Appendix 1. □

Proposition 8 suggests that under the social security system, the optimal intergenerational contract, $\widetilde{\delta}^k$, is governed by four parameters, namely $(\beta, r, \varepsilon_h, \varepsilon_l)$. To further demonstrate some properties of $\widetilde{\delta}^k$, we develop the next proposition:

Proposition 9

(i) $\frac{d\widetilde{\delta}^k}{d\alpha} = 0$;
(ii) $\frac{d\widetilde{\delta}^k}{d\beta} < 0$ *if* $\ln\frac{\beta\delta}{1+r} > 2$.

Proof See Appendix 1. □

It is clear from (25) that the optimal intergenerational contract is independent of the preference for old-age consumption (i.e., $\widetilde{\delta}^k$ is not a function of α). We explain Proposition 9(i) in two aspects. First, given the Cobb-Douglas utility function in (24), the young-age consumption is proportional to the old-age consumption in equilibrium. Second, as Proposition 2 shows, parents' investments in children's schooling are not affected by their consumption preferences. Proposition 9(ii) demonstrates that $\widetilde{\delta}^k$ decreases with β under certain conditions. Intuitively, as the improved educational efficiency increases the income level, individuals become sufficiently rich to feed themselves and less reliant on children's provision of old-age support.

When social security is absent, we write the steady-state social welfare in (11) as $U^n(\delta)$. Therefore, the optimal intergenerational intrafamily trade, $\widetilde{\delta}^n$, is a function of $(\alpha, \beta, r, \varepsilon_l, \varepsilon_h)$, although it is mathematically difficult to find an analytical solution. We hence rely on a simulation to show the properties of $\widetilde{\delta}^n$ and then compare $\widetilde{\delta}^n$ with $\widetilde{\delta}^k$. We aim to simulate the comparative statics of $\widetilde{\delta}$ with respect to $(\alpha, \beta, r, \varphi(w))$, where $\varphi(w)$ is a function of $(\varepsilon_l, \varepsilon_h)$ by Eqs. (2)–(4):

$$\varphi(w) := \frac{\sqrt{E(w^2)-[E(w)]^2}}{E(w)} = \frac{H(e)\sqrt{E[(1+\varepsilon)^2]-1}}{H(e)E[(1+\varepsilon)]} = \sqrt{E(\varepsilon^2)} = \sqrt{-\varepsilon_l\varepsilon_h}. \tag{26}$$

Interpreted literally, $\varphi(w)$ represents the coefficient of variation of the steady-state wage rate.

To run the simulation, we use parameter values as in Table 1. In particular, the intertemporal interest rate equals 60%, which is equivalent to a return of deposits

Table 1 Benchmark values of parameters in simulation

Parameter	Interpretation	Value
α	Preference for old-age consumption	3
β	Efficiency of education	100
r	Intertemporal interest rate	60%
ε_h	Good luck in the job market	0.4
ε_l	Bad luck in the job market	− 0.4

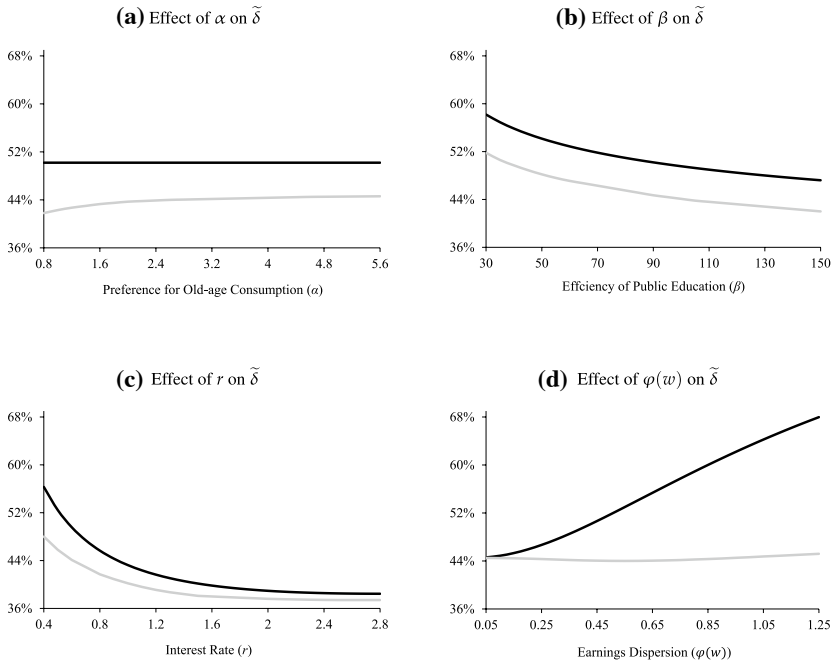

Fig. 2 The optimal intergenerational contract. *Note*: The dark curve represents the case with social security, and the gray curve represents the case without social security

at 2% compounded annually for a period of 24 years. Figure 2 charts four sets of simulation results, where dark and gray curves represent the cases with and without social security, respectively. In all plots, the dark curve lies above the gray curve, showing that the optimal transfer is smaller without social security. This is because parents face uncertainty of their children's types and cannot know how much they will receive from children in future. Hence, they count less on children's wealth transfer but preserve a greater disposable income and also save more.

Plot (a) charts the relationship between the preference for old-age consumption and the optimal intergenerational contract. The dark curve is horizontal, which reflects the result from Proposition 9(i) that $\widetilde{\delta}^{k}$ is unrelated to α under the social security system. However, the gray curve exhibits a rising trend. Holding other things constant, an increased marginal utility from old-age consumption motivates people to substitute old-age for young-age consumption. To accomplish this, parents will design a rule requiring their children to afford a larger transfer; moreover, this rule is binding on themselves, thereby leading to a cut in their young-age disposable income and consumption.

Plot (b) shows that $\widetilde{\delta}$ decreases with β no matter whether social security is present. As we have analyzed in Proposition 9(ii), the simple intuition is that good education and high income in young adulthood make retirees less dependent on family financial support. In plot (c), the interest rate is negatively related to the fraction of income transferred to parents. Intuitively, the increased returns to savings not only add to individuals' lifetime income but also prompt them to save more, making intergenerational transfers less necessary.

Plot (d) graphs the optimal intergenerational contract given various degrees of the relative variability of a generation's earnings. We illustrate the values of $\varphi(w)$ by altering the wage premium ε_h and holding $\varepsilon_l = -0.4$. The dark curve slopes upward: greater earnings dispersion results in a more generous social security. However, the gray curve looks flatter, and it first follows a downtrend and then a upward trend. Specifically, the optimal fraction of earnings transferred to parents without social security decreases from 44.50% at $\varphi(w) = 0.05$ to 44.01% at $\varphi(w) = 0.55$ and then rises steadily to 45.19% at $\varphi(w) = 1.25$.

6 Conclusion and further discussions

The analysis of risk sharing is central to public economics in general and social protection in particular (Boadway and Keen 2000; Fleurbaey 2018). This paper builds a model with heterogeneous agents analyzing the role of social security in sharing intragenerational labor market risk through intergenerational links. Based on insurance principles, social security enables individuals to effectively hedge against the risk of their children's earnings. It hence serves as a substitute for the traditional kinship practice to pool risks of retirement benefits across families.

In our model, children are *ex ante* identical obtaining the same public education but become *ex post* different in earning capacities due to the job market uncertainty. Parents invest in education because it increases their children's future productivities and labor earnings. In return, they receive children's wealth transfers in old age in accordance with either an implicit intrafamily contract or a PAYGO social security scheme. If parents count on their own children for old-age financial support, then educational investments are risky because children's earnings are uncertain. The replacement of traditional kinship practices with social security transforms investments in children's education from a risky to safe asset: under the social security scheme, what parents can get in old age depends on children's average earnings, for

which there is no idiosyncratic risk for any parent. As human capital becomes less risky to invest in, more educational investments are induced, and all children are thus better educated. Our analysis suggests that the PAYGO social security scheme leads to an improvement in risk management of the parents' generation and an increase in human capital accumulation of the children's generation, which creates welfare gains.

We also suggest that under the public school system, the social security scheme worsens the economic inequality in steady state. If intergenerational transfer occurs within a family, a low-paid individual gets a considerable transfer if her child becomes a high earner, while a high-paid individual gets a small transfer if her child has a low earning capacity. In other words, an individual's lifetime income depends on not only her own wage but also her child's wage. Because of the uncertainty of children's future earnings, individuals move up or down the income ladder throughout the lifetime, which helps to curb wealth polarization. However, social security effectively equalizes retirement benefits received by the whole generation, which removes the intragenerational mobility and hence makes the income distribution less equal.

We then examine the optimal intergenerational contract, which is the fraction of earnings to be transferred that maximizes the steady-state social welfare. We consider comparative statics for the model parameters and find that the optimal transfer under the social security scheme is independent of people's preferences over consumption in old age versus in young age. Our simulation exercises suggest that the optimal transfer to the old is generally higher under the social security scheme.

Following Ehrlich and Lui (1991), Chakrabarti et al. (1993) and Rangel (2003), we assume that children are a pure investment good (i.e., a child's earning capacity does not enter her parent's utility function) in our model. This assumption helps us to sharpen our main message, while consideration of parental altruism (Barro 1974; Becker 1974) is not incompatible. Our results continue to hold qualitatively if we introduce both the child's ascending transfer (*investment motive*) and his human capital (*altruistic motive*) into the parent's utility function. Besides, as Cox and Stark (2005) suggest, an individual who cares about her grandchild is willing to transfer wealth to her child by, for example, providing housing downpayments for her grandchild's living space and paying his tuition fees. But if her child gets a high-paying job by virtue of good education, he would be afford housing and schooling by himself without accepting her pecuniary aid. In this aspect, our analysis is close in spirit to the Barro-Becker altruism model.

Our model assumes that an individual does not obtain utility from her child's human capital and earnings, which further leads us to focus only on public education. In future research, we may follow the literature (e.g., Becker and Gregg Lewis 1973; Galor and Weil 2000; Fan and Zhang 2013) to incorporate the child's quality into the parent's utility function. In this case, it is expected that the rich will spend more on children's education, and hence, their children are likely to earn a higher wage in future. Consequently, the result of intragenerational equality obtained in our paper will to some extent be modified. However, it will not affect the qualitative result of our paper that poor parents may experience some mobility due to the

uncertainties that are associated with their children's future earnings in the absence of social security.

With a focus on the risk-mitigating effect of social security, our model considers that all parents have one child. Nonetheless, the literature on endogenous fertility holds that fertility rate is often negatively related to income (e.g., de la Croix and Doepke 2004; Cremer et al. 2006). When the poor have more children, it will be welfare improving to reallocate educational resources from rich children to poor children by the law of diminishing returns. A practical way of this reallocation is to implement a public educational system in which each child receives the same educational resources, as analyzed in our paper. As children grow up and transfer part of their incomes to their old parents, the whole old generation tends to benefit from the early provision of better education for poor children. Therefore, the consideration of endogenous fertility would help to reinforce the main result of our model.

A large literature on dynamic efficiency has developed since Samuelson's (1958) seminal contribution, which addresses the determination of the efficient direction of intergenerational transfers by comparing population growth with interest rates. This literature holds that in an infinite-horizon economy, the PAYGO social security that allocates resources from the young to the old in each period is Pareto-improving if the population growth rate exceeds the interest rate (e.g., Abel et al. 1989). Our analysis complements this literature by suggesting that social security makes both the young and the old better off even if the population growth rate falls below the interest rate. The channel for social security to improve welfare is that it insures people against the uncertainty of retirement benefits, which fosters economic development by encouraging human capital investments.

Appendix 1: Mathematical proofs

Proof of Lemma 1

Substituting (7) into (6) yields

$$U_{it} = u[(1-\delta)(1+\varepsilon_i)H(e_{t-1}) - s_{it} - e_t)] + v[(1+r)s_{it} - f_t], \tag{27}$$

where f_t is independent of s_{it} by (8). The maximization problem of an i-type member implies

$$u'[(1-\delta)(1+\varepsilon_i)H(e_{t-1}) - s_{it} - e_t)] = (1+r)v'[(1+r)s_{it} - f_t]. \tag{28}$$

Suppose that $s_{lt} \geqslant s_{ht}$. It follows directly from (28) that

$$\begin{aligned} & u'[(1-\delta)(1+\varepsilon_h)H(e_{t-1}) - s_{ht} - e_t)] \\ & \quad = (1+r)v'[(1+r)s_{ht} - f_t] \\ & \quad \geqslant (1+r)v'[(1+r)s_{lt} - f_t] \\ & \quad = u'[(1-\delta)(1+\varepsilon_l)H(e_{t-1}) - s_{lt} - e_t)] \\ & \quad > u'[(1-\delta)(1+\varepsilon_h)H(e_{t-1}) - s_{ht} - e_t)], \end{aligned} \tag{29}$$

where the first inequality results from $v'' < 0$, and the second one results from $u'' < 0$ and $\varepsilon_l < \varepsilon_h$. Because (29) is a self-contradiction, the presumption $s_{lt} \geqslant s_{ht}$ violates, which means $s_{lt} < s_{ht}$.

Proof of Proposition 1

Expanding the first-order condition (12) obtains

$$\begin{aligned}
\frac{dU_t^n}{de_t^n} &= pu'[(1-\delta)(1+\varepsilon_h)H(e_{t-1}^n) - s_{ht}^n - e_t^n]\left(-\frac{ds_{ht}^n}{de_t^n} - 1\right) \\
&\quad + p^2 v'[(1+r)s_{ht}^n + \delta(1+\varepsilon_h)H(e_t^n)] \\
&\quad \left[(1+r)\frac{ds_{ht}^n}{de_t^n} + \delta(1+\varepsilon_h)H'(e_t^n)\right] \\
&\quad + p(1-p)v'[(1+r)s_{ht}^n + \delta(1+\varepsilon_l)H(e_t^n)] \\
&\quad \left[(1+r)\frac{ds_{ht}^n}{de_t^n} + \delta(1+\varepsilon_l)H'(e_t^n)\right] \\
&\quad + (1-p)u'[(1-\delta)(1+\varepsilon_l)H(e_{t-1}^n) - s_{lt}^n - e_t^n]\left(-\frac{ds_{lt}^n}{de_t^n} - 1\right) \\
&\quad + p(1-p)v'[(1+r)s_{lt}^n + \delta(1+\varepsilon_h)H(e_t^n)] \\
&\quad \left[(1+r)\frac{ds_{lt}^n}{de_t^n} + \delta(1+\varepsilon_h)H'(e_t^n)\right] \\
&\quad + (1-p)^2 v'[(1+r)s_{lt}^n + \delta(1+\varepsilon_l)H(e_t^n)] \\
&\quad \left[(1+r)\frac{ds_{lt}^n}{de_t^n} + \delta(1+\varepsilon_l)H'(e_t^n)\right] = 0 \\
\Leftrightarrow & p^2 v'[(1+r)s_{ht}^n + \delta(1+\varepsilon_h)H(e_t^n)]\left[\delta(1+\varepsilon_h)H'(e_t^n) - (1+r)\right] \\
&\quad + p(1-p)v'[(1+r)s_{ht}^n + \delta(1+\varepsilon_l)H(e_t^n)] \\
&\quad \left[\delta(1+\varepsilon_l)H'(e_t^n) - (1+r)\right] \\
&\quad + p(1-p)v'[(1+r)s_{lt}^n + \delta(1+\varepsilon_h)H(e_t^n)] \\
&\quad \left[\delta(1+\varepsilon_h)H'(e_t^n) - (1+r)\right] \\
&\quad + (1-p)^2 v'[(1+r)s_{lt}^n + \delta(1+\varepsilon_l)H(e_t^n)] \\
&\quad \left[\delta(1+\varepsilon_l)H'(e_t^n) - (1+r)\right] = 0 \\
\Leftrightarrow & \Phi_t p\left[(1+\varepsilon_h)H'(e_t^n) - \frac{1+r}{\delta}\right] \\
&\quad + \Omega_t(1-p)\left[(1+\varepsilon_l)H'(e_t^n) - \frac{1+r}{\delta}\right] = 0,
\end{aligned} \tag{30}$$

where two notations are introduced:

$$\begin{aligned}\Phi_t &:= pv'[(1+r)s^n_{ht} + \delta(1+\varepsilon_h)H(e^n_t)] \\ &\quad + (1-p)v'[(1+r)s^n_{lt} + \delta(1+\varepsilon_h)H(e^n_t)] > 0,\end{aligned} \tag{31}$$

$$\begin{aligned}\Omega_t &:= pv'[(1+r)s^n_{ht} + \delta(1+\varepsilon_l)H(e^n_t)] \\ &\quad + (1-p)v'[(1+r)s^n_{lt} + \delta(1+\varepsilon_l)H(e^n_t)] > 0.\end{aligned} \tag{32}$$

By $\varepsilon_h > \varepsilon_l$ and $H' > 0$, we can infer from (30)–(32) that

$$\frac{1+r}{\delta(1+\varepsilon_h)} < H'(e^n_t) < \frac{1+r}{\delta(1+\varepsilon_l)}. \tag{33}$$

Since (10) presents s^n_{it} as a function of (e^n_{t-1}, e^n_t), (30) implies that e^n_t is time-varying for some t. Because $v'' < 0$ and $\varepsilon_h > \varepsilon_l$, we have $\Phi_t < \Omega_t$. Rewrite (30) and then use (4) to obtain

$$\begin{aligned}&[\Phi_t p(1+\varepsilon_h) + \Omega_t(1-p)(1+\varepsilon_l)]H'(e^n_t) = \frac{1+r}{\delta}[\Phi_t p + \Omega_t(1-p)] \\ &\Leftrightarrow \frac{1+r}{\delta H'(e_t)} - 1 = \frac{\Phi_t p\varepsilon_h + \Omega_t(1-p)\varepsilon_l}{\Phi_t p + \Omega_t(1-p)} < \frac{\Omega_t[p\varepsilon_h + (1-p)\varepsilon_l]}{\Phi_t p + \Omega_t(1-p)} = 0 \\ &\Leftrightarrow H'(e^n_t) > \frac{1+r}{\delta}.\end{aligned} \tag{34}$$

Combining (33) and (34) derives Eq. (13).

Proof of Proposition 2

Taking the first-order condition of (16) with respect to e^k_t and then inserting (15) obtains

$$
\begin{aligned}
\frac{dU_t^k}{de_t^k} = p\Bigg\{ & -u'[(1-\delta)(1+\varepsilon_h)H(e_{t-1}^k) - s_{ht}^k - e_t^k]\left(1+\frac{\partial s_{ht}^k}{\partial e_t^k}\right) \\
& + v'[(1+r)s_{ht}^k + \delta H(e_t^k)]\left[(1+r)\frac{\partial s_{ht}^k}{\partial e_t^k} + \delta H'(e_t^k)\right]\Bigg\} + (1-p) \\
& \Bigg\{ -u'[(1-\delta)(1+\varepsilon_l)H(e_{t-1}^k) - s_{lt}^k - e_t^k]\left(1+\frac{\partial s_{lt}^k}{\partial e_t^k}\right) \\
& + v'\left[(1+r)s_{lt}^k + \delta H(e_t^k)\right]\left[(1+r)\frac{\partial s_{lt}^k}{\partial e_t^k} + \delta H'(e_t^k)\right]\Bigg\} = 0 \\
\Leftrightarrow & pu'[(1-\delta)(1+\varepsilon_h)H(e_{t-1}^k) - s_{ht}^k - e_t^k] \\
& + (1-p)u'[(1-\delta)(1+\varepsilon_l)H(e_{t-1}^k) - s_{lt}^k - e_t^k] \\
& \left\{pv'[(1+r)s_{ht}^k + \delta H(e_t^k)] + (1-p)v'[(1+r)s_{lt}^k + \delta H(e_t^k)]\right\} \\
& \delta H'(e_t^k) = 0 \\
\Leftrightarrow & \left[H'(e_t^k) - \frac{1+r}{\delta}\right] \\
& \left\{\frac{pv'[(1+r)s_{ht}^k + \delta H(e_t^k)]}{1-p} + v'[(1+r)s_{lt}^k + \delta H(e_t^k)]\right\} = 0.
\end{aligned} \tag{35}
$$

Since $p \in (0, 1)$ and $v' > 0$, it is straightforward to infer from (35) that (17) holds.

Proof of Proposition 3

The coefficient of variation of f_t, as denoted by $\varphi(f_t)$, equals the standard deviation of f_t to the mean of f_t, namely,

$$
\varphi(f_t) := \frac{SD(f_t)}{E(f_t)} = \frac{\sqrt{E(f_t^2) - [E(f_t)]^2}}{E(f_t)}, \tag{36}
$$

Without social security, we can use (4) and (8) to rewrite (36) as

$$
\varphi(f_t^n) = \frac{\sqrt{[\delta H(e_t)]^2[p\varepsilon_h^2 + (1-p)\varepsilon_l^2]}}{\delta H(e_t)} = \sqrt{\frac{\varepsilon_l\varepsilon_h^2}{\varepsilon_l - \varepsilon_h} - \frac{\varepsilon_l^2\varepsilon_h}{\varepsilon_l - \varepsilon_h}} = \sqrt{-\varepsilon_l\varepsilon_h}. \tag{37}
$$

With social security, we can use (8) to rewrite (36) as

$$
\varphi(f_t^k) = \frac{\sqrt{[\delta H(e_t)]^2 - [\delta H(e_t)]^2}}{\delta H(e_t)} = 0. \tag{38}
$$

Proof of Proposition 4

Propositions 1 and 2 imply that $H'(e^n) > H'(e^k)$. Therefore, $e^n < e^k$ under the assumption of $H'' < 0$.

Proof of Proposition 5

We can infer from (17) and (19) that $H'(e^*) < H'(e^k)$ for $\delta \in (0, 1)$. By the assumption $H'' < 0$ and Proposition 4, we have $e^* > e_t^k > e_t^n$.

Proof of Proposition 6

Given that $v'' < 0$, for any w_{it}, s_{it}, and e_t where $i \in \{h, l\}$, we can infer from Eq. (11):

$$\begin{aligned}
U^n(e_t) &= p\big\{u[(1-\delta)w_{ht} - s_{ht} - e_t] + Ev[(1+r)s_{ht} + \delta(1+\varepsilon)H(e_t)]\big\} \\
&\quad + (1-p)\big\{u[(1-\delta)w_{lt} - s_{lt} - e_t] + Ev[(1+r)s_{lt} + \delta(1+\varepsilon)H(e_t)]\big\} \\
&< p\big\{u[(1-\delta)w_{ht} - s_{ht} - e_t] + v[(1+r)s_{ht} + \delta E(1+\varepsilon)H(e_t)]\big\} \\
&\quad + (1-p)\big\{u[(1-\delta)w_{lt} - s_{lt} - e_t] + v[(1+r)s_{lt} + \delta E(1+\varepsilon)H(e_t)]\big\} \\
&= p\big\{u[(1-\delta)w_{ht} - s_{ht} - e_t] + v[(1+r)s_{ht} + \delta H(e_t)]\big\} \\
&\quad + (1-p)\big\{u[(1-\delta)w_{lt} - s_{lt} - e_t] + v[(1+r)s_{lt} + \delta H(e_t)]\big\} = U^k(e_t).
\end{aligned} \tag{39}$$

It follows that for any e_t^n, $\max\{U_t^k\} \geqslant U^k(e_t^n) > U^n(e_t^n) = \max\{U_t^n\}$.

Proof of Proposition 7

We first examine the steady-state Gini coefficient without social security. By (21), the total income of each generation amounts to $H(e^n)$. Referring to Figure 1(a), we obtain λ_1 and λ_3 as follows:

$$\lambda_1 = \frac{(1-p)^2[(1+\varepsilon_l)H(e^n)]}{H(e^n)} = (1-p)^2(1+\varepsilon_l), \tag{40}$$

$$\begin{aligned}
\lambda_3 &= \lambda_1 + \frac{p(1-p)[1+(1-\delta)\varepsilon_l + \delta\varepsilon_h]H(e^n) + p(1-p)[1+(1-\delta)\varepsilon_h + \delta\varepsilon_l]H(e^n)}{H(e^n)} \\
&= (1-p)[1+(1+\varepsilon_l)p].
\end{aligned} \tag{41}$$

However, the expression of λ_2 depends on the income ranking. In case that $\delta < \frac{1}{2}$, the ranking is $I_{ll} < I_{lh} < I_{hl} < I_{hh}$. It follows that

$$\lambda_2 = \lambda_1 + \frac{p(1-p)[1+(1-\delta)\varepsilon_l + \delta\varepsilon_h]H(e^n)}{H(e^n)} = (1-p)[1+(1-\delta)\varepsilon_l]. \tag{42}$$

In case that $\delta \geqslant \frac{1}{2}$, the income ranking is $I_{ll} < I_{hl} \leqslant I_{lh} < I_{hh}$. It follows that

$$\lambda_2 = \lambda_1 + \frac{(1-p)p[1+(1-\delta)\varepsilon_h + \delta\varepsilon_l]H(e^n)}{H(e^n)} = (1-p)(1+\delta\varepsilon_l). \tag{43}$$

Plugging (40)–(43) into (22) and rearranging obtains G^n as

$$G^n = \begin{cases} (1-p)[2p(1-p)\delta\varepsilon_l - \varepsilon_l] & \text{if } \delta < \frac{1}{2} \\ (1-p)[2p(1-p)(1-\delta)\varepsilon_l - \varepsilon_l] & \text{otherwise} \end{cases}, \tag{44}$$

We then examine the Gini coefficient with social security. The share of total income earned by all l-type members can be expressed by

$$\lambda_4 = \frac{(1-p)(1+\varepsilon_l)H(e^k)}{(1-p)(1+\varepsilon_l)H(e^k) + p(1+\varepsilon_h)H(e^k)} = (1-p)(1+\varepsilon_l). \tag{45}$$

Inserting (45) into (23) yields $G^k = -(1-p)\varepsilon_l$. By this equation and (44), we show that

$$G^n - G^k = \begin{cases} 2p(1-p)^2\delta\varepsilon_l < 0 & \text{if } \delta < \frac{1}{2} \\ 2p(1-p)^2(1-\delta)\varepsilon_l < 0 & \text{otherwise} \end{cases}. \tag{46}$$

Proof of Proposition 8

Given the specific functions in (24), the optimal educational spending per child can be written as

$$e_t^k = \frac{\beta\delta}{1+r}. \tag{47}$$

Plugging (47) into (15) solves the i-type individual's optimal (interior) saving level as

$$\begin{aligned} &\frac{1}{\beta(1-\delta)(1+\varepsilon_i)\ln e_{t-1}^k - s_{it}^k - e_t} = \frac{\alpha(1+r)}{(1+r)s_{it}^k + \beta\delta\ln e_t^k} \\ &\Leftrightarrow (1+\alpha)s_{it}^k = \alpha\left[\beta(1-\delta)(1+\varepsilon_i)\ln\frac{\beta\delta}{1+r} - \frac{\beta\delta}{1+r}\right] - \frac{\beta\delta}{1+r}\ln\frac{\beta\delta}{1+r}, \\ &\Leftrightarrow s_{it}^k = \frac{\beta}{1+\alpha}\left[\alpha(1-\delta)(1+\varepsilon_i) - \frac{\delta}{1+r}\right]\ln\frac{\beta\delta}{1+r} - \frac{\alpha\beta\delta}{(1+\alpha)(1+r)} \end{aligned} \tag{48}$$

Substituting (47) and (48) into (16) derives the steady-state social welfare under the social security system as

$$
\begin{aligned}
U^k = p \ln & \\
& \left[\frac{\beta(1-\delta)(1+\varepsilon_h)}{1+\alpha}\ln\frac{\beta\delta}{1+r} + \frac{\beta\delta}{(1+\alpha)(1+r)}\ln\frac{\beta\delta}{1+r} - \frac{\beta\delta}{(1+\alpha)(1+r)}\right] \\
& + \alpha p \ln\left[\frac{\alpha\beta(1+\delta)(1+\varepsilon_h)(1+r)}{1+\alpha}\ln\frac{\beta\delta}{1+r} + \frac{\alpha\beta\delta}{1+\alpha}\ln\frac{\beta\delta}{1+r} - \frac{\alpha\beta\delta}{1+r}\right] \\
& + (1-p)\ln\left[\frac{\beta(1-\delta)(1+\varepsilon_l)}{1+\alpha}\ln\frac{\beta\delta}{1+r} + \frac{\beta\delta}{(1+\alpha)(1+r)}\ln\frac{\beta\delta}{1+r} - \frac{\beta\delta}{(1+\alpha)(1+r)}\right] \\
& + \alpha(1-p)\ln\left[\frac{\alpha\beta(1+\delta)(1+\varepsilon_l)(1+r)}{1+\alpha}\ln\frac{\beta\delta}{1+r} + \frac{\alpha\beta\delta}{1+\alpha}\ln\frac{\beta\delta}{1+r} - \frac{\alpha\beta\delta}{1+r}\right] \\
= \ln & \left[\frac{\alpha^\alpha}{1+r}\left(\frac{\beta}{1+\alpha}\right)^{1+\alpha}\right] \\
& + p(1+\alpha)\ln\left\{[(1+r)(1+\varepsilon_h)(1-\delta)+\delta]\ln\frac{\beta\delta}{1+r} - \delta\right\} \\
& + (1-p)(1+\alpha)\ln\left\{[(1+r)(1+\varepsilon_l)(1-\delta)+\delta]\ln\frac{\beta\delta}{1+r} - \delta\right\}.
\end{aligned} \tag{49}
$$

Differentiating (49) with respect to δ and setting it to zero yields the interior solution to δ:

$$
\begin{aligned}
\frac{\partial U_t^k}{\partial\delta} &= p(1+\alpha)\left\{\frac{-(r+\varepsilon_h+r\varepsilon_h)\ln\frac{\beta\delta}{1+r} + [(1+r)(1+\varepsilon_h)(1-\delta)+\delta]/\delta - 1}{[(1+r)(1+\varepsilon_h)(1-\delta)+\delta]\ln\frac{\beta\delta}{1+r} - \delta}\right\} \\
&\quad + (1-p)(1+\alpha)\left\{\frac{-(r+\varepsilon_l+r\varepsilon_l)\ln\frac{\beta\delta}{1+r} + [(1+r)(1+\varepsilon_l)(1-\delta)+\delta]/\delta - 1}{[(1+r)(1+\varepsilon_l)(1-\delta)+\delta]\ln\frac{\beta\delta}{1+r} - \delta}\right\} = 0 \\
&\Leftrightarrow \frac{1-p}{p}\left\{\frac{(1+r+\varepsilon_l+r\varepsilon_l)(1-\delta) - \delta(r+\varepsilon_l+r\varepsilon_l)\ln\frac{\beta\delta}{1+r}}{[(1+r+\varepsilon_l+r\varepsilon_l)(1-\delta)+\delta]\ln\frac{\beta\delta}{1+r} - \delta}\right\} \\
&\quad + \frac{(1+r+\varepsilon_h+r\varepsilon_h)(1-\delta) - \delta(r+\varepsilon_h+r\varepsilon_h)\ln\frac{\beta\delta}{1+r}}{[(1+r+\varepsilon_h+r\varepsilon_h)(1-\delta)+\delta]\ln\frac{\beta\delta}{1+r} - \delta} = 0 \\
&\Leftrightarrow \frac{\varepsilon_h}{\varepsilon_l}\frac{(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right) + 1 - \delta}{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r} - \delta} \\
&\quad = \frac{(r+\varepsilon_h+r\varepsilon_h)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right) + 1 - \delta}{[1+(r+\varepsilon_h+r\varepsilon_h)(1-\delta)]\ln\frac{\beta\delta}{1+r} - \delta},
\end{aligned}
$$

which is equivalent to Eq. (25).

Proof of Proposition 9

It is clear from (25) that $\widetilde{\delta}^k$ is not a function of α. Besides, totally differentiating (25) with respect to δ and β obtains:

$$
\begin{aligned}
&\frac{\left\{(r+\varepsilon_h+r\varepsilon_h)\left[-d\delta-\ln\frac{\beta\delta}{1+r}d\delta-\delta\left(\frac{d\beta}{\beta}+\frac{d\delta}{\delta}\right)\right]-d\delta\right\}\left[(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta\right]}{\left[(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta\right]^2}\\
&\quad-\frac{\left\{(r+\varepsilon_l+r\varepsilon_l)\left[-d\delta-\ln\frac{\beta\delta}{1+r}d\delta-\delta\left(\frac{d\beta}{\beta}+\frac{d\delta}{\delta}\right)\right]-d\delta\right\}\left[(r+\varepsilon_h+r\varepsilon_h)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta\right]}{\left[(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta\right]^2}\\
&=\frac{\left\{-(r+\varepsilon_h+r\varepsilon_h)\ln\frac{\beta\delta}{1+r}d\delta+[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\left(\frac{d\beta}{\beta}+\frac{d\delta}{\delta}\right)-d\delta\right\}\left\{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta\right\}}{\frac{\varepsilon_l}{\varepsilon_h}\left\{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta\right\}^2}\\
&\quad-\frac{\left\{-(r+\varepsilon_l+r\varepsilon_l)\ln\frac{\beta\delta}{1+r}d\delta+[1+(r+\varepsilon_h+r\varepsilon_h)(1-\delta)]\left(\frac{d\beta}{\beta}+\frac{d\delta}{\delta}\right)-d\delta\right\}\left\{[1+(r+\varepsilon_h+r\varepsilon_h)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta\right\}}{\frac{\varepsilon_l}{\varepsilon_h}\left\{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta\right\}^2}\\
&\Leftrightarrow\frac{(\varepsilon_h+r\varepsilon_h-\varepsilon_l-r\varepsilon_l)(1-\delta)\left(-2d\delta-\ln\frac{\beta\delta}{1+r}d\delta-\frac{\delta}{\beta}d\beta\right)+(\varepsilon_h+r\varepsilon_h-\varepsilon_l-r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)d\delta}{\left[(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta\right]^2}\\
&=\frac{(\varepsilon_h+r\varepsilon_h-\varepsilon_l-r\varepsilon_l)\left\{\left[(1-\delta)\frac{d\beta}{\beta}+(1-\delta)\frac{d\delta}{\delta}-\ln\frac{\beta\delta}{1+r}d\delta\right]\left(\ln\frac{\beta\delta}{1+r}-\delta\right)+(1-\delta)\ln\frac{\beta\delta}{1+r}\left(d\delta-\frac{d\beta}{\beta}-\frac{d\delta}{\delta}\right)\right\}}{\frac{\varepsilon_l}{\varepsilon_h}\left\{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta\right\}^2}\\
&\Leftrightarrow\frac{\frac{\delta}{\beta}d\beta+\left(1+\frac{1}{1-\delta}\ln\frac{\beta\delta}{1+r}\right)d\delta}{\left[(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta\right]^2}\\
&=\frac{\frac{\delta}{\beta}d\beta+\left(1+\frac{1}{1-\delta}\ln\frac{\beta\delta}{1+r}\right)d\delta+\frac{1}{1-\delta}\left(\ln\frac{\beta\delta}{1+r}\right)^2d\delta-\frac{2}{1-\delta}\ln\frac{\beta\delta}{1+r}d\delta}{\frac{\varepsilon_l}{\varepsilon_h}\left\{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta\right\}^2}\\
&\Leftrightarrow\frac{\varepsilon_l}{\varepsilon_h}\left\{\frac{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta}{(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta}\right\}^2-1\\
&=\frac{\frac{1}{1-\delta}\ln\frac{\beta\delta}{1+r}\left(\ln\frac{\beta\delta}{1+r}-2\right)}{\left(1+\frac{1}{1-\delta}\ln\frac{\beta\delta}{1+r}\right)+\frac{\delta}{\beta}\frac{d\beta}{d\delta}}\\
&\Leftrightarrow\frac{d\beta}{d\delta}=\frac{\beta}{\delta}\left\{\frac{\frac{1}{1-\delta}\ln\frac{\beta\delta}{1+r}\left(\ln\frac{\beta\delta}{1+r}-2\right)}{\frac{\varepsilon_l}{\varepsilon_h}\left\{\frac{[1+(r+\varepsilon_l+r\varepsilon_l)(1-\delta)]\ln\frac{\beta\delta}{1+r}-\delta}{(r+\varepsilon_l+r\varepsilon_l)\left(1-\delta-\delta\ln\frac{\beta\delta}{1+r}\right)+1-\delta}\right\}^2-1}-\left(1+\frac{1}{1-\delta}\ln\frac{\beta\delta}{1+r}\right)\right\},
\end{aligned}
$$

where the first term in the bracket of the right-hand-side is negative if $\ln\frac{\beta\delta}{1+r}-2>0$, while the second term is positive. Therefore, $\frac{d\delta}{d\beta}<0$ if $\ln\frac{\beta\delta}{1+r}>2$.

Appendix 2: A numerical example of Sect. 3.1

This appendix provides a simulation on the evolution of public educational expenditure, private savings, and social welfare without social security. Our numerical exercise rests on specific functional forms in (24). We first rewrite the i-type individual's maximization problem (10) as a quadratic function of s_{it}^n:

Table 2 Transitional dynamics and steady state under the intrafamily transfer scheme

Variable	Period 0	Period 1	Period 2	Period 3	Period 4	...	Steady State
Educational spending per child (e^n)	9	17.076	17.497	17.511	17.511	...	17.511
Private saving of the poor (s_l^n)	–	24.907	38.552	39.076	39.092	...	39.093
Private saving of the rich (s_h^n)	–	122.441	165.031	166.656	166.708	...	166.709
Social welfare (U^n)	–	19.970	20.775	20.803	20.804	...	20.804

$$
\begin{aligned}
&\frac{\alpha(1+r)p}{(1+r)s_{it}^n+\beta\delta(1+\varepsilon_h)\ln(e_t^n)}+\frac{\alpha(1+r)(1-p)}{(1+r)s_{it}^n+\beta\delta(1+\varepsilon_l)\ln(e_t^n)}\\
&\quad=\frac{1}{\beta(1-\delta)(1+\varepsilon_i)\ln(e_{t-1}^n)-s_{it}^n-e_t^n},\\
&\Leftrightarrow \alpha\left[s_{it}^n+\frac{\beta\delta}{1+r}(1+\varepsilon_h+\varepsilon_l)\ln(e_t^n)\right]=\frac{\left[s_{it}^n+\frac{\beta\delta(1+\varepsilon_h)\ln(e_t^n)}{1+r}\right]\left[s_{it}^n+\frac{\beta\delta(1+\varepsilon_l)\ln(e_t^n)}{1+r}\right]}{\beta(1-\delta)(1+\varepsilon_i)\ln(e_{t-1}^n)-s_{it}^n-e_t^n}\\
&\Leftrightarrow (s_{it}^n)^2+\left\{\frac{\beta\delta\ln e_t^n}{1+r}\left(\varepsilon_h+\varepsilon_l+\frac{2+\alpha}{1+\alpha}\right)-\frac{\alpha[\beta(1-\delta)(1+\varepsilon_i)\ln e_{t-1}^n-e_t^n]}{1+\alpha}\right\}s_{it}^n\\
&\quad+\left(\frac{\beta\delta\ln e_t^n}{1+r}\right)^2\frac{(1+\varepsilon_h)(1+\varepsilon_l)}{1+\alpha}\\
&\quad-\frac{\alpha\beta\delta(1+\varepsilon_h+\varepsilon_l)[\beta(1-\delta)(1+\varepsilon_i)\ln e_{t-1}^n-e_t^n]\ln e_t^n}{(1+\alpha)(1+r)}=0.
\end{aligned}
\tag{50}
$$

The government's optimal choice of e_t^n characterized by (12) can be solved in the next equation:

$$
\begin{aligned}
&\frac{p[(1+r)(ps_{lt}^n+(1-p)s_{ht}^n)+\beta\delta(1+\varepsilon_h)\ln e_t^n]\left[(1+r)e_t^n-\beta\delta(1+\varepsilon_h)\right]}{[(1+r)s_{ht}^n+\beta\delta(1+\varepsilon_h)\ln e_t^n][(1+r)s_{lt}^n+\beta\delta(1+\varepsilon_h)\ln e_t^n]}\\
&\quad=\frac{(p-1)[(1+r)(ps_{lt}^n+(1-p)s_{ht}^n)+\beta\delta(1+\varepsilon_l)\ln e_t^n]\left[(1+r)e_t^n-\beta\delta(1+\varepsilon_l)\right]}{[(1+r)s_{ht}^n+\beta\delta(1+\varepsilon_l)\ln e_t^n][(1+r)s_{lt}^n+\beta\delta(1+\varepsilon_l)\ln e_t^n]},
\end{aligned}
\tag{51}
$$

where $p=\frac{\varepsilon_l}{\varepsilon_l-\varepsilon_h}$ by (4). Combining (50) and (51) implies that e_t^n is governed by e_{t-1}^n and six parameters, namely $(e_{t-1}^n;\alpha,\beta,r,\varepsilon_h,\varepsilon_l,\delta)$.

To do the simulation, we rely on the values of parameters as in Table 1. In addition, we set the initial level of educational spending to 9 (i.e., $e_0^n=9$) and the fraction of earnings transferred to parents to 0.4 (i.e., $\delta=40\%$). Table 2 reports the simulated results of e_t^n, s_{lt}^n, s_{ht}^n, and U_t^n in every period t. For example, starting from 9 in the period 0, educational spending per child stands at 17.076 in period 1, increases to 17.497 in period 2, and then converges to its steady-state level of 17.511. The other three variables exhibit a similar pattern: private savings and social welfare rise gradually to their respective steady-state levels.

Acknowledgements We are grateful to the anonymous referee for valuable comments and suggestions. We also benefit from helpful comments of conference participants of Society for the Advancement of Economic Theory and Association of Public Economic Theory in 2019. Yu Pang gratefully acknowledges financial support provided by Macau University of Science and Technology Foundation. All errors are our own.

References

Abel, A., Mankiw, G., Summers, L., & Zeckhauser, R. (1989). Assessing dynamic efficiency: Theory and evidence. *Review of Economic Studies, 56*(1), 1–19.

Arteaga, C. (2018). The effect of human capital on earnings: Evidence from a reform at Colombia's top university. *Journal of Public Economics, 157,* 212–225.

Baland, J.-M., & Robinson, J. (2000). Is child labor inefficient? *Journal of Political Economy, 108*(4), 663–679.

Ball, L., & Mankiw, G. (2007). Intergenerational risk sharing in the spirit of Arrow, Debreu, and Rawls, with applications to social security design. *Journal of Political Economy, 115*(4), 523–547.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy, 82*(6), 1095–1117.

Becker, G. S. (1974). A theory of social interactions. *Journal of Political Economy, 82*(6), 1063–1093.

Becker, G. S. (1993). Nobel lecture: The economic way of looking at behavior. *Journal of Political Economy, 101*(3), 385–409.

Becker, G. S., & Gregg Lewis, H. (1973). On the interaction between the quantity and quality of children. *Journal of Political Economy, 81*(2), S279–S288.

Bellettini, G., & Ceroni, C. B. (2000). Social security expenditure and economic growth: An empirical assessment. *Research in Economics, 54*(3), 249–275.

Bloom, D., Canning, D., Mansfield, R., & Moore, M. (2007). Demographic change, social security systems, and savings. *Journal of Monetary Economics, 54*(1), 92–114.

Boadway, R. (2006). Intergovernmental redistributive transfers: Efficiency and equity. *Handbook of fiscal federalism, Chapter 14* (pp. 355–380). Cheltenham: Edward Elgar.

Boadway, R., & Keen, M. (2000). Redistribution. In A. B. Atkinson & F. Bourguignon (Eds.), *Handbook of income distribution, Chapter 12* (Vol. 1, pp. 677–789). New York: Elsevier.

Boadway, R., & Marchand, M. (1995). The use of public expenditures for redistributive purposes. *Oxford Economic Papers, 47*(1), 45–59.

Bowles, S., Gintis, H., & Osborne, M. (2001). The determinants of earnings: A behavioral approach. *Journal of Economic Literature, 39*(4), 1137–1176.

Bozio, A., Emmerson, C., O'Dea, C., & Tetlow, G. (2017). Do the rich save more? Evidence from linked survey and administrative data. *Oxford Economic Papers, 69*(4), 1101–1119.

Brown, C. (1980). Equalizing differences in the labor market. *Quarterly Journal of Economics, 94*(1), 113–134.

Chakrabarti, S., Lord, W., & Rangazas, P. (1993). Uncertain altruism and investment in children. *American Economic Review, 83*(4), 994–1002.

Cheung, S. N. S. (1972). The enforcement of property rights in children, and the marriage contract. *Economic Journal, 82*(326), 641–657.

Cox, D., & Stark, O. (2005). On the demand for grandchildren: Tied transfers and the demonstration effect. *Journal of Public Economics, 89*(9–10), 1665–1697.

Cremer, H., Gahvari, F., & Pestieau, P. (2006). Pensions with endogenous and stochastic fertility. *Journal of Public Economics, 90*(12), 2303–2321.

Cremer, H., Gahvari, F., & Pestieau, P. (2011). Fertility, human capital accumulation, and the pension system. *Journal of Public Economics, 95*(11–12), 1272–1279.

Cremer, H., Lozachmeur, J.-M., & Pestieau, P. (2007). Social security and retirement decision: A positive and normative approach. *Journal of Economic Surveys, 22*(2), 213–233.

Cremer, H., & Roeder, K. (2017). Social insurance with competitive insurance markets and risk misperception. *Journal of Public Economics, 146*(C), 138–147.

de la Croix, D., & Doepke, M. (2004). Public versus private education when differential fertility matters. *Journal of Development Economics, 73*(2), 607–629.

de la Croix, D., & Doepke, M. (2009). To segregate or to integrate: Education politics and democracy. *Review of Economic Studies, 76*(2), 597–628.

de la Croix, D., & Michel, P. (2002). *A theory of economic growth: Dynamics and policy in overlapping generations*. New York: Cambridge University Press.

Deaton, A., Gourinchas, P.-O., & Paxson, C. (2002). Social security and inequality over the life cycle. In M. Feldstein & J. Liebman (Eds.), *The distributional aspects of social security and social security reform* (pp. 115–148). Chicago: University of Chicago Press.

Diamond, P. (1977). A framework for social security analysis. *Journal of Public Economics, 8*(3), 275–298.

Drèze, J., & Khera, R. (2017). Recent social security initiatives in India. *World Development, 98,* 555–572.

Dynan, K., Skinner, J., & Zeldes, S. (2004). Do the rich save more? *Journal of Political Economy, 112*(2), 397–444.

Echevarría, C. A., & Iza, A. (2006). Life expectancy, human capital, social security and growth. *Journal of Public Economics, 90*(12), 2323–2349.

Ehrlich, I., & Kim, J. (2007). Social security and demographic trends: Theory and evidence from the international experience. *Review of Economic Dynamics, 10*(1), 55–77.

Ehrlich, I., & Lui, F. T. (1991). Intergenerational trade, longevity, and economic growth. *Journal of Political Economy, 99*(5), 1029–1059.

Ehrlich, I., & Lui, F. T. (1998). Social security, the family, and economic growth. *Economic Inquiry, 36*(3), 390–409.

Enders, W., & Lapan, H. E. (1982). Social security taxation and intergenerational risk sharing. *International Economic Review, 23*(3), 647–658.

Fafchamps, M., & Lund, S. (2003). Risk-sharing networks in rural Philippines. *Journal of Development Economics, 71*(2), 261–287.

Fan, S., & Zhang, J. (2013). Differential fertility and intergenerational mobility under private versus public education. *Journal of Population Economics, 26*(3), 907–941.

Fleurbaey, M. (2008). *Fairness, responsibility, and welfare*. Oxford: Oxford University Press.

Fleurbaey, M. (2018). Welfare economics, risk and uncertainty. *Canadian Journal of Economics, 51*(1), 5–40.

Fuster, L., İmrohoroğlu, A., & İmrohoroğlu, S. (2003). A welfare analysis of social security in a dynastic framework. *International Economic Review, 44*(4), 1247–1274.

Galor, O., & Weil, D. (2000). Population, technology, and growth: From Malthusian stag-nation to the demographic transition and beyond. *American Economic Review, 90*(4), 806–828.

Gandelman, N. (2017). Do the rich save more in Latin America? *Journal of Economic Inequality, 15*(1), 75–92.

Glomm, G., & Kaganovich, M. (2008). Social security, public education and the growth-inequality relationship. *European Economic Review, 52*(6), 1009–1034.

Glomm, G., & Ravikumar, B. (1992). Public versus private investment in human capital: Endogenous growth and income inequality. *Journal of Political Economy, 100*(4), 818–834.

Gokhale, J., Kotlikoff, L. J., Sefton, J., & Weale, M. (2001). Simulating the transmission of wealth inequality via bequests. *Journal of Public Economics, 79*(1), 93–128.

Gordon, R. H., & Varian, H. R. (1988). Intergenerational risk sharing. *Journal of Public Economics, 37*(2), 185–202.

Hartog, J., van Ophem, H., & Bajdechi, S. M. (2007). Simulating the risk of investment in human capital. *Education Economics, 15*(3), 259–275.

Krueger, D., & Kubler, F. (2002). Intergenerational risk-sharing via social security when financial markets are incomplete. *American Economic Review, 92*(2), 407–410.

Lee, C.-C., & Chang, C.-P. (2006). Social security expenditure and GDP in OECD countries: A cointegrated panel analysis. *International Economic Journal, 20*(3), 303–320.

Ljungqvist, L., & Sargent, T. J. (2014). Career length: Effects of curvature of earnings profiles, earnings shocks, taxes, and social security. *Review of Economic Dynamics, 17*(1), 1–20.

Lucas, R., Jr. (1992). On efficiency and distribution. *Economic Journal, 102*(411), 233–247.

Luttmer, E., & Samwick, A. (2018). The welfare cost of perceived policy uncertainty: Evidence from social security. *American Economic Review, 108*(2), 275–307.

Merton, R. C. (1983). On the role of social security as a means for efficient risk sharing in an economy where human capital is not tradable. In Z. Bodie & J. Shoven (Eds.), *Financial aspects of the United States pension system* (pp. 325–358). Chicago: University of Chicago Press.

Nishiyama, S., & Smetters, K. (2007). Does social security privatization produce efficiency gains? *Quarterly Journal of Economics, 122*(4), 1677–1719.
Organisation for Economic Cooperation and Development (OECD). (2020). General Government Spending (Indicator). https://doi.org/10.1787/a31cbf4d-en. Accessed 5 February 2020.
Pestieau, P., & Lefebvre, M. (2018). *The welfare state in Europe: Economic and social perspectives* (2nd ed.). Oxford: Oxford University Press.
Pestieau, P., & Possen, U. M. (1979). A model of wealth distribution. *Econometrica, 47*(3), 761–772.
Pluchino, A., Biondo, A., & Rapisarda, A. (2019). Exploring the role of talent and luck in getting success. *Acta Physica Polonica B Proceedings Supplement, 12*(1), 17–24.
Rangel, A. (2003). Forward and backward intergenerational goods: Why is social security good for the environment? *American Economic Review, 93*(3), 813–834.
Samuelson, P. (1958). An exact consumption-loan model of interest with or without the social contrivance of money. *Journal of Political Economy, 66*(6), 467–482.
Shea, J. (2000). Does parents' money matter? *Journal of Public Economics, 77*(2), 155–184.
Shiller, R. J. (1999). Social security and institutions for intergenerational, intragenerational, and international risk-sharing. *Carnegie-Rochester Conference Series on Public Policy, 50,* 165–204.
Shiller, R. J. (2003). Social security and individual accounts as elements of overall risk-sharing. *American Economic Review, 93*(2), 343–347.
Sinn, H.-W. (2004). The pay-as-you-go pension system as a fertility insurance and enforcement device. *Journal of Public Economics, 88*(7–8), 1335–1357.
Storesletten, K., Telmer, C. I., & Yaron, A. (1999). The risk-sharing implications of alternative social security arrangements. *Carnegie-Rochester Conference Series on Public Policy., 50,* 213–259.
Townsend, R. (1994). Risk and insurance in village India. *Econometrica, 62*(3), 539–591.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.