**Simon Fan, Yu Pang and
Pierre Pestieau**

A model of the optimal allocation of government expenditures

CORE

CENTER FOR
OPERATIONS RESEARCH
AND ECONOMETRICS

A Model of the Optimal Allocation of Government Expenditures

Simon Fan   Yu Pang   Pierre Pestieau[*]

Abstract

Government expenditures can be used for various socio-economic objectives, including public education, consumption of public goods and services, and social protection. This paper analyzes the optimal allocation of public expenditures among these competing functions. We establish an overlapping generations model with heterogeneous individuals in which the government optimally chooses income tax, transfer payment, educational spending, and public consumption. Our model characterizes the transitional dynamics and the steady state of each function with and without a pay-as-you-go intergenerational contract. We also conduct a simulation illustrating that the presence of an intergenerational contract may raise public consumption and social welfare in the steady state.

JEL Classifications: H20, H31, H50

Keywords: Government Spending; Public Education; Public Consumption; Individual Heterogeneity

1   Introduction

Large-scale public expenditures are a widespread phenomenon in industrialized economies. In 2016, general government spending amounted to 37.8% of the gross domestic product in the U.S., 39% in Japan, and about 50% in many European countries (OECD 2018). Government's massive involvement in market economies serves redistributive, allocative, and regulatory purposes, which effectively mitigate income inequality, provide public goods and services, and promote sustainable development (Musgrave 1959, Barr 1992). For instance, the Nordic model, which features a blend of socialist values and free market capitalism, is usually perceived to have succeeded in advancing equity, efficiency, and well-being (Stiglitz 2015, Pestieau and Lefebvre 2018).

Government spending has a variety of socio-economic objectives. Table 1 lists the government expenditure breakdown by ten functions, and reports their percentage shares of total expenditures in several European countries between 2012 and 2016. In each country, the largest share is devoted to social protection, mainly in the form of income-redistribution programs that transfer wealth to the poor. Schooling is also an important target of public funds: most of these governments appropriate around 10% of their expenditures to establish and maintain the state education systems. The other eight functions, including health care, public services, economic affairs (e.g., infrastructure construction and supply of water and electricity), environmental protection, etc., fit generally into the category of consumption of public goods and services. Public consumption accounts for half of government spending in almost every country.

Based on the classifications used in Table 1, we categorize government expenditures into three main divisions – public education, social protection, and consumption of public goods and services. Educational expenditure provides two benefits: first, it facilitates the formation of human capital of the future generation, allowing altruistic parents to benefit from their children's improved quality; second, it promotes social mobility, which helps lower wage inequality of future generations. The provision of public goods and services directly enhances people's welfare. As economic progress improves the living standard, households want more public consumption, which is complementary to their private consumption (Ganelli and Tervala 2009). Transfers to the poor not only raise social welfare by the law of diminishing marginal utility (e.g., Boadway and Keen 2000) but also address the public concern over income inequality and long-term economic development (De la Croix and

Table 1. Government Expenditure by Function (% of Total Expenditure): 2012–2016 Average

	Belgium	Denmark	France	Germany	Greece	Italy	Netherlands	Spain	UK
Social protection	36.3	43.3	43.0	42.8	37.6	41.7	36.5	38.9	37.9
Education	11.6	12.6	9.6	9.6	8.2	7.9	11.8	9.1	11.6
Health	14.2	15.5	14.2	15.9	9.3	14.1	17.8	13.7	17.3
General public services	15.4	13.7	11.5	14.1	18.6	17.3	10.9	14.9	11.4
Economic affairs	12.7	6.4	9.4	7.2	13.7	8.0	8.9	11.3	7.2
Public order & safety	3.3	1.8	2.9	3.5	3.8	3.8	4.2	4.5	4.6
Defence	1.6	2.2	3.1	2.4	4.4	2.4	2.5	2.1	5.0
Environment protection	1.8	0.8	1.8	1.4	2.7	1.8	3.3	1.9	1.8
Recreation, culture & religion	2.3	3.2	2.4	2.2	1.3	1.5	3.2	2.6	1.7
Community amenities	0.7	0.5	2.1	0.9	0.5	1.4	0.9	1.1	1.6

Data source: *Eurostat* (accessed on July 25, 2018). Government expenditures are decomposed into ten broad functions based on the classification of the functions of government (COFOG) used in the European regime of National Accounts (ESA 2010). Figures are calculated as a simple average based on the available data for the most recent five years.

Doepke 2003, 2004, Fleurbaey 2008, 2009).

To the extent that various categories compete for resources under the fiscal budget constraint, an intriguing question arises as to how government spending can be optimally allocated between them? The analysis of the optimal allocation of government spending has deep roots in economics, going back at least to Samuelson (1954). Existing literature on public finance concentrates on one or two categories. For example, Docquier, Paddison, and Pestieau (2007) and Del Rey and Lopez-Garcia (2013) study the optimal levels of education and pension provided by a welfare-maximizing government; other studies include Possen and Slutsky (1980), Meltzer and Richard (1981), Glomm and Ravikumar (1992, 1997), Boadway and Marchand (1995), Kaganovich and Zilcha (1999), Del Rey and Lopez-Garcia (2016). This paper develops a more comprehensive profile covering social protection, public education, public consumption, and intergenerational mobility.

We establish an overlapping generations model in which the government chooses income taxes/transfers and expenditures on education and public consumption. In our model, there are two types of individuals – the skilled and the unskilled – whose income levels are positively related to their labor skill. Addressing individual heterogeneity allows us to probe into the redistributive function of government spending. Moreover, an individual's skill plays a vital role in shaping her children's

quality, as a number of empirical studies have demonstrated that parental background is critical to an individual's achievement (e.g., Becker 1991, De la Croix and Michel 2002). Therefore, even if individuals' net incomes tend to be equalized after taxes/transfers, the effect of skill heterogeneity on children's human capital formation persists.

We start our analysis with the case in which each generation care about their own utility and do not make any wealth transfer to the previous generation. Our analysis shows that in the steady state educational expenditure increases with the learning efficiency of skilled workers' children and the wage differential but decreases with the fertility rate. Also, under some configuration, the expenditure on public consumption is inversely related to the learning efficiency of skilled workers' children but positively related to the wage differential and the fertility rate.

We next turn to the case where generations are linked through a social contract in which each generation contribute a certain fraction of their incomes to public consumption of the previous generation.[1] The logic of the intergenerational contract resembles a pay-as-you-go (PAYG) program, which requires that workers pay social security taxes into the fund and money flows back out to retirees. We examine the transitional dynamics of different categories of government expenditures, showing how the introduction of an intergenerational contract changes the results. We also discuss the intergenerational contract that maximizes social welfare in the steady state.

To compare the two cases, we conduct a numerical simulation illustrating the transitional dynamics and the steady states of educational expenditure, public consumption, and social welfare. We find that the optimal contract allows each generation to consume more public goods and services than their parents did till the steady-state level is achieved. In contrast, the absence of ascending transfers may result in regressive public consumption over time, largely because of the crowding-out effect of the expanding educational expenditure. Under certain parameter configurations, the presence of the optimal contract will be welfare-improving.

The remainder of this paper is organized as below. Section 2 lays out our model setup. Section 3 analyzes the optimal allocation of government spending when every generation care about their own utility. Section 4 considers an intergenerational contract under which the young generation financially supports public consumption of the parents' generation. Section 5 performs a quantita-

---

[1]The concept "generational accounting", which was developed in the 1990s (Auerbach, Gokhale, and Kotlikoff 1994), is considered an important aspect of government expenditures in national income accounting (Mason and Tapinos 2000, Lee and Mason 2011).

3

tive analysis comparing the results of sections 3 and 4. Section 6 examines the robustness of the model. Section 7 concludes. Most mathematical proofs are relegated to the Appendix.

2 The Model Setup

Consider an overlapping generations economy where individuals live for three periods: childhood, young adulthood, and old adulthood. In childhood, an individual receives a publicly funded education but does not make decisions. In young adulthood, each individual earns a wage, chooses her private consumption and saving levels, and rears n children. In old adulthood, all individuals retire, spend their savings on private consumption without leaving bequests, and consume public goods provided by the government.[2]

Individuals who spend the young adulthood in period t are labeled as "Generation t". A fraction $\theta_t \in (0, 1)$ of generation t become skilled workers (s), while the others serve as unskilled workers (u). Preferences of each i-type member where $i \in \{s, u\}$ are given by

$$v_t^i = \ln c_{yt}^i + \alpha \ln c_{ot}^i + \beta q_t^i + \phi \ln G_t, \tag{2.1}$$

where c_{yt}^i and c_{ot}^i denote the private consumption of an i-type member of generation t in her young and old adulthood, q_t^i denotes the quality of her children, G_t represents her public consumption in old adulthood, and α, β, and ϕ are positive parameters. Note that G_t is the same for all members of generation t. In the beginning of period t, a representative member of generation t expects to obtain v_t^s with a probability θ_t and v_t^u with a probability $1 - \theta_t$. In line with Harsanyi (1955), the social welfare is defined as the following weighted sum:

$$V_t = \theta_t v_t^s + (1 - \theta_t) v_t^u. \tag{2.2}$$

The quality of an i-type member's child is measured by the probability that she will grow up to be a skilled worker, which depends on her family background (indicated by i) and the government's

---

[2]To highlight the essence of the model, we assume that individuals consume public goods in old adulthood only. In reality, seniors need more environmental protection and health care because they are more at risk from environmental threats and are susceptible to age-related chronic diseases and decreased physical strength (Geller and Zenick 2005). This assumption serves the purpose of algebraic simplification. Our main results continue to hold if we assume that public goods are consumed throughout the life cycle.

educational expenditure per child, E_t:

$$q_t^i = \begin{cases} \ln(1 + \sigma E_t) & \text{if } i = u \\ \rho \ln(1 + \sigma E_t) & \text{if } i = s \end{cases}, \tag{2.3}$$

where $q_t^i \in (0, 1)$. The parameter $\rho > 1$ measures the degree of elitism of public education (e.g., Cremer, De Donder and Pestieau 2010). Holding the educational expenditure constant, a child whose parent is skilled is more likely to grow up to be skilled (i.e., $q^s > q^u$).[3] The parameter $\sigma > 0$ indicates that the increased educational expenditure improves the quality of *any* child (i.e., q^i increases with E). To ensure the probability $q^i < 1$ requires

$$\rho \ln(1 + \sigma E_t) < 1 \quad \Rightarrow \quad E_t < \frac{e^{\frac{1}{\rho}} - 1}{\sigma}, \tag{2.4}$$

where e is Euler's number. For simplicity, we assume that (2.4) holds throughout this paper.

If N_t denotes the population of generation t and n denotes the fertility rate, the population of generation $t + 1$ (i.e., the number of children in period t) amounts to

$$N_{t+1} = nN_t. \tag{2.5}$$

Given that there are $\theta_t N_t$ skilled workers and $(1 - \theta_t)N_t$ unskilled workers in generation t, they will bear $q_t^s(n\theta_t N_t)$ and $q_t^u[n(1 - \theta_t)N_t]$ skilled children, respectively. With the total number of skilled workers in the next generation being $q_t^s n\theta_t N_t + q_t^u n(1 - \theta_t)N_t$, the proportion of skilled workers in generation $t + 1$ can be written as

$$\theta_{t+1} = \frac{q_t^s n\theta_t N_t + q_t^u n(1 - \theta_t)N_t}{N_{t+1}} = [1 + (\rho - 1)\theta_t]\ln(1 + \sigma E_t), \tag{2.6}$$

which captures the evolution of θ_t.

We examine the production of the economy to determine the wage of workers by using Galor and Zeira's (1993) modeling strategy. The economy consists of a skilled labor and an unskilled labor sector, both producing the unique final good competitively. Production in the unskilled labor

---

[3]Different from the existing literature on social mobility (e.g., Cremer, De Donder, and Pestieau 2010), we assume that children's education does not depend on parental educational decisions.

sector requires unskilled labor only. Without loss of generality, one unit of unskilled labor can be converted into one unit of output:

$$Y^u = L^u, \tag{2.7}$$

where Y^u and L^u are the output and labor input of the unskilled labor sector, respectively. With the price of a final good being normalized to unity, unskilled workers obtain an equilibrium wage of one in a perfectly competitive market. In the skilled labor sector, the output Y^s is produced by capital K and skilled labor L^s according to a constant-return-to-scale production function F:

$$Y^s = F(K, L^s) = L^s f(k), \qquad \text{where } k \equiv K/L^s, \tag{2.8}$$

where k denotes the ratio of capital to skilled labor. If both factor markets are competitive, we derive the interest rate r, and the wage rate of skilled labor w as

$$r = f'(k), \qquad w = f(k) - kf'(k). \tag{2.9}$$

In a small open economy with free capital mobility, the interest rate equals the world's (exogenous) interest rate $r = \bar{r}$. Since k remains fixed at $f'^{-1}(r)$, w is constant and governed by r. In line with empirical findings (e.g., Broecke 2016), we assume that a worker's (pre-tax) wage rate is positively related to her skill ($w > 1$). Note that w also represents the wage differential between skilled and unskilled workers, i.e., the ratio of a skilled worker's wage to an unskilled worker's wage.

The government can identify the sector in which an individual works and impose a lump-sum income tax from workers.[4] Denote T_t^i as the payment of each i-type member of generation t to the government, which is a tax if positive, and a social transfer if negative. Since T^u can be negative, the government can freely redistribute income from the skilled to the unskilled. Accordingly, the government obtains a total revenue of $N_t \theta_t T_t^s + N_t(1 - \theta_t)T_t^u$ from generation t.

On the expenditure side, the government spends $N_{t+1}E_t$ on children's education and defrays the elderly's public consumption. We consider two channels of sponsoring public consumption. Section 3 examines public consumption as pre-funded by generation t themselves. The government

---

[4]Brito, Hamilton, Slutsky, and Stiglitz (1991) and Hamilton and Slutsky (2007) posit that government can identify individuals' types and assess personalized lump-sum taxes. Our paper makes the same assumption. For the case where types are not observable, see Cremer, Pestieau, and Rochet (2001).

appropriates a present value of $\frac{N_t G_t}{1+r}$ in period t for public consumption so that each member of generation t consumes G_t in period $t+1$ when she ages. Section 4 analyzes an intergenerational social contract, which requires generation $t+1$ to contribute a fixed fraction of their income to support the public consumption of generation t.

3 No Intergenerational Contract

This section investigates the case in which every generation care about their own utility but not the previous generation's utility. Although a parent regards her children's educational attainments as a consumption good (equation (2.1)), a child has no incentive to enhance his parent's welfare. In other words, altruism is assumed to be descending only.

In period t, the interaction between the government and generation t proceeds in two stages. The government moves first to maximize the social welfare of generation t by choosing educational expenditure per child (E_t), public consumption per member of generation t in her old adulthood (G_t), and taxation and/or transfer (T_t^s, T_t^u). After observing government policies, each i-type member of generation t optimally chooses her private consumption (c_{yt}^i, c_{ot}^i) where $i \in \{s, u\}$.

By backward induction, we first determine an individual's decisions given government policies. In period t, the lifetime budget constraints of an i-type worker can be expressed by

$$c_{yt}^i + \frac{c_{ot}^i}{1+r} = I_t^i, \tag{3.1}$$

where I_t^i is an i-type member's after-tax income net of child-rearing expenses in period t:

$$I_t^i = \begin{cases} 1 - T_t^u - \delta n & \text{if } i = u \\ w - T_t^s - \delta n & \text{if } i = s \end{cases}, \tag{3.2}$$

where δ denotes the cost of rearing a child. Each i-type member of generation t chooses her private consumptions to maximize her utility (2.1) subject to her budget constraint (3.1), which solves

$$c_{yt}^i = \frac{I_t^i}{1+\alpha}, \qquad c_{ot}^i = \frac{\alpha(1+r)I_t^i}{1+\alpha}. \tag{3.3}$$

In period t, the government collects taxes (T_t^s, T_t^u) from skilled and unskilled workers, spends

7

part of the revenue on public education (E_t), and saves money for generation t's public consumption in their old adulthood (G_t). The government's budget constraint requires that total spending not be more than total tax revenues:

$$N_{t+1}E_t + \frac{N_t G_t}{1+r} \leqslant \theta_t N_t T_t^s + (1-\theta_t)N_t T_t^u \quad \Rightarrow \quad nE_t + \frac{G_t}{1+r} \leqslant \theta_t T_t^s + (1-\theta_t)T_t^u. \quad (3.4)$$

Facing this budget constraint, the government aims to maximize the social welfare of generation t, which can be derived by inserting (2.1) and (3.3) into (2.2):

$$V_t = \Delta + (1+\alpha)\theta_t \ln(w - \delta n - T_t^s) + (1+\alpha)(1-\theta_t)\ln(1 - \delta n - T_t^u)$$
$$+ \beta[1 + (\rho - 1)\theta_t]\ln(1 + \sigma E_t) + \phi \ln G_t, \quad (3.5)$$

where $\Delta := \alpha \ln[\alpha(1+r)] - (1+\alpha)\ln(1+\alpha)$ is a constant.

To solve the constrained maximization problem, we write the Lagrangian as

$$\mathbb{L}_t = \Delta + (1+\alpha)\theta_t \ln(w - \delta n - T_t^s) + (1+\alpha)(1-\theta_t)\ln(1 - \delta n - T_t^u) + \phi \ln G_t$$
$$+ \beta[1 + (\rho - 1)\theta_t]\ln(1 + \sigma E_t) - \lambda_t \left\{ nE_t + \frac{G_t}{1+r} - \theta_t T_t^s - (1-\theta_t)T_t^u \right\}, \quad (3.6)$$

where λ_t represents the Lagrange multiplier associated with the government revenue constraint. Taking the first order conditions with respect to E_t, G_t, T_t^s, T_t^u, and λ_t and then rearranging solves the government's optimal policies as

$$E_t = \frac{\beta[1 + (\rho - 1)\theta_t][1 + (w - 1)\theta_t - \delta n]/n - (1 + \alpha + \phi)/\sigma}{1 + \alpha + \beta + \phi + \beta(\rho - 1)\theta_t}, \quad (3.7)$$

$$G_t = \frac{\phi(1+r)[1 + (w-1)\theta_t + (1/\sigma - \delta)n]}{1 + \alpha + \beta + \phi + \beta(\rho - 1)\theta_t}, \quad (3.8)$$

$$T_t^s = w - \frac{(1+\alpha)[1 + (w-1)\theta_t + (1/\sigma - \delta)n]}{1 + \alpha + \beta + \phi + \beta(\rho - 1)\theta_t}, \quad (3.9)$$

$$T_t^u = 1 - \frac{(1+\alpha)[1 + (w-1)\theta_t + (1/\sigma - \delta)n]}{1 + \alpha + \beta + \phi + \beta(\rho - 1)\theta_t}. \quad (3.10)$$

Based on equations (3.7) and (3.8), we demonstrate the relationship between θ_t and (E_t, G_t) in the following proposition:

Proposition 1 *In period t, a larger proportion of skilled workers leads to*

(a) greater expenditure on public education per child ($\frac{\partial E_t}{\partial \theta_t} > 0$);

(b) lower expenditure on public consumption ($\frac{\partial G_t}{\partial \theta_t} < 0$) if and only if

$$\frac{\rho - 1}{w - 1}\left[1 + \left(\frac{1}{\sigma} - \delta\right)n\right] \geqslant \frac{1 + \alpha + \phi}{\beta} + 1. \tag{3.11}$$

Proof. See Appendix. ∎

Proposition 1(a) suggests that all else being equal, a government tends to spend more on public education as the proportion of skilled workers (i.e., the rich) in the working-age population increases. This result is consistent with the observation that, in general, developed countries spend a greater percentage of national income on public education than developing countries (World Bank 2017). Intuitively, when a larger proportion of people become rich, they want to invest more in education to improve the quality of their children, which is a normal good of consumption (equation (2.1)). This reflects the typical "income effect." In addition, given that rich children have higher learning efficiency than poor children, the larger the proportion of rich children, the higher the return on schooling and the more the educational spending ("family background effect"). From the perspective of fiscal budgeting, when an increased number of rich people pay more taxes, the government's budget increases and the educational spending increases accordingly ("fiscal budget effect"). Moreover, as we can see from equation (2.6), a larger proportion of skilled workers, coupled with an increase in contemporaneous educational expenditure, expands the future proportion of skilled workers, which in turn boosts economic development.

Proposition 1(b) shows that G_t and θ_t have an ambiguous relationship. In particular, public consumption may decrease with the proportion of skilled workers if and only if condition (3.11) is satisfied. Although the "income effect" and "fiscal budget effect" remain applicable in Proposition 1(b), the substitution (resources competition) between public education and public consumption potentially lowers the level of public consumption. This substitution effect become more prominent when people have a weak preference for public goods (small ϕ) and the children of skilled workers have improved learning efficiency (large ρ). Also, if an improvement in children's quality gives rise to significant gains (large β) but educational investments generate low returns for children (small σ), then the government needs to increase its expenditure on public education, which

9

in turn crowds out its expenditure on public consumption.

Since the expenditures on public goods are formulated as a component of the log-linear utility function, one may expect that it always increases with economic development, as measured by the proportion of skilled workers. However, Proposition 1(b) shows a somewhat surprising result that it is not always the case. The intriguing relationship between the expenditures on public goods and economic development depends critically on condition (3.11).

In Proposition 1, we discuss a trait of (E_t, G_t) in transitional dynamics. Following the earlier studies (Mookherjee and Napel 2007, Fan and Stark 2008, Cremer, Gahvari, and Pestieau 2011), we define that labor force composition dynamics reach a steady state when the proportion of skilled workers remain unchanging over time:

Definition 1 *The labor force composition dynamics is in the steady state if the proportion of skilled workers is constant, i.e., $\theta_t = \theta$ for all t.*

Note that the proportion of skilled workers in the working-age population is an essential indicator of economic development. Based on equation (2.6) and Definition 1, we derive

$$\theta = [1 + (\rho - 1)\theta] \ln(1 + \sigma E) \qquad \Rightarrow \qquad E = \frac{e^{\frac{\theta}{1+(\rho-1)\theta}} - 1}{\sigma}. \qquad (3.12)$$

The next proposition follows directly from (3.7) and (3.12):

Proposition 2 *The steady-state proportion of skilled workers, θ, is determined by*

$$\frac{\theta}{1 + (\rho - 1)\theta} = \ln \left\{ \frac{\beta[1 + (\rho - 1)\theta][1 - \delta\sigma + \sigma/n + (w - 1)\theta\sigma/n]}{1 + \alpha + \beta + \phi + \beta(\rho - 1)\theta} \right\}. \qquad (3.13)$$

The steady state always exists, which is nontrivial and stable.

Proof. See Appendix. ∎

Proposition 2 characterizes the steady-state proportion of skilled workers in the young population. Equation (3.13) implies that θ is a function of eight parameters, namely, $(\alpha, \beta, \phi, \delta, \rho, \sigma, w, n)$. The solution to (3.13), which falls within (0,1), always exists, and it is a stable equilibrium. In the long run, the proportion of skilled workers remain constant across different generations.

By (3.12) and (3.13), we develop the following proposition:

Proposition 3 *In the steady state, the proportion of skilled workers and the educational spending*

per child

(a) increase with the learning efficiency differential ($\frac{d\theta}{d\rho} > 0$ and $\frac{dE}{d\rho} > 0$);

(b) decrease with the fertility rate ($\frac{d\theta}{dn} < 0$ and $\frac{dE}{dn} < 0$);

(c) increase with the wage differential ($\frac{d\theta}{dw} > 0$ and $\frac{dE}{dw} > 0$).

Proof. See Appendix. ∎

Proposition 3(a) is intuitive: holding other things unchanged, if children of skilled workers have improved learning efficiency (larger ρ), then the government tends to spend more on education (larger E), and all children are less likely to become unskilled workers (larger θ). Proposition 3(b) demonstrates that all else being equal, as people rear fewer children (smaller n), each child receives more educational spending (larger E), which leads to the improved quality of children. This result is in line with Becker and Lewis's (1973) theorem that parents face a tradeoff between children's quantity and quality if children are treated as pure consumption goods. While Becker and Lewis (1973) and the subsequent research focus on the individual household's decision on the number of children, our analysis differs in two aspects. The quality of children depends on the educational expenditure chosen by the government, which aims to maximize the utilitarian welfare on behalf of all parents. Also, we assume that the fertility rate is exogenous for simplicity. In other words, Proposition 3(b) presents a normative analysis of the optimal choice of the quality of children in response to different levels of fertility from the perspective of social welfare.

According to Proposition 3(c), educational expenditure increases with skilled workers' income holding unskilled workers' income constant. This result captures the "income effect": when people become richer, they spend more in promoting children's quality, which is a normal good in their utility function. Moreover, since all children have access to the same educational resources, an increase in skilled workers' earnings will enhance the educational attainment of all children, despite that family background continues to play a crucial role in children's human capital formation.

We establish the following proposition to examine the comparative static of the steady-state expenditure on public consumption:

Proposition 4 *The steady-state expenditure on public consumption per person*

(a) decreases with the learning efficiency differential ($\frac{\partial G}{\partial \rho} < 0$) if condition (3.11) holds;

11

(b) increases with the fertility rate $\left(\frac{\partial G}{\partial n} > 0\right)$ *if condition (3.11) holds and*

$$\delta\sigma < 1; \tag{3.14}$$

(c) increases with the wage differential $\left(\frac{\partial G}{\partial w} > 0\right)$ *if condition (3.11) is not satisfied.*

Proof. See Appendix. ■

Proposition 4 reveals that the public consumption level critically depends on condition (3.11), which concerns the substitution effect between E and G. Proposition 4(a) indicates that the expenditure on public consumption may be inversely related to the learning efficiency of skilled workers' children. Specifically, when public consumption and public education are sufficiently substitutable, increased educational spending induced by the improved learning efficiency of skilled workers' children crowds out the expenditure on public consumption. Next, we find a positive effect of fertility rate on public consumption under some configurations. Proposition 4(b) implies that given a lower population growth rate, people tend to sacrifice their old-age consumption of public goods if they have a weak preference for public goods (small ϕ) and a strong preference for children's quality (large β), but the returns on educational investment are generally low (small σ). Proposition 4(c) holds that a higher wage of skilled workers may result in a higher level of public consumption per person. This positive relationship can be explained by "income effect" and "fiscal budget effect", and it is more likely to be true when the substitution effect between E and G is dominated.

We next turn to the features of income tax and transfer. The next proposition follows directly from equations (3.9) and (3.10):

Proposition 5 *The net income after taxation and subsidy is equal for skilled and unskilled workers.*

Proposition 5 shows that the lifetime net incomes (i.e., incomes after taxation and redistribution) of skilled and unskilled workers equalize in each period, i.e., $w - T^s = 1 - T^u$. This result sheds light on the facts shown in Table 1 that social protection accounts for about 40 percent of all government expenditures in many European countries. Proposition 5 follows from our assumption that income taxation does not entail administrative cost or have any distortionary effect on labor supply. Note that despite the same after-tax income, a skilled individual obtains a higher utility than an unskilled individual in that her children are of higher quality.

The following proposition investigates the effects of ρ, n, and w on (T^s, T^u):

Proposition 6 *In the steady state,*

(a) T^s and T^u increase with the learning efficiency differential ρ if condition (3.11) holds;

(b) T^s and T^u increase with the fertility rate n if condition (3.11) and (3.14) hold;

(c) T^s increases with the wage differential w if condition (3.11) holds, and T^u decreases with the wage differential w if condition (3.11) is not satisfied.

Proof. See Appendix. ∎

Proposition 6(a) suggests that when family background plays a more salient role in children's skill development, the government may set a higher income tax (or a lower transfer payment if any). Intriguingly, we find that the poor may pay more taxes or receive less transfer payments when rich children become even more efficient in human capital formation. This is because the improved learning efficiency of rich children leads to a greater proportion of the rich in the steady state (Proposition 3(a)), which raises the average income level. If people obtain a high utility from improved quality of children (large β), the government will collect more taxes so as to allocate more resources to public education. Although the poor may have a lower net income, their children become better educated because of the government's greater effort in education.

According to Proposition 6(b), to the extent that the population grows faster (larger n), the working-age population will have to contribute more taxes to finance total educational spending. This is likely to be true when people gain more from the improvement in their children's quality (large β) and it is not too costly to rear children (small δ). Proposition 6(c) posits that the relationship between income tax and wage differential depends on condition (3.11), which is a determinant of government expenditures.

4   Intergenerational Social Contract

In this section, we consider the case in which young adults support their parents' public consumption through direct wealth transfers. It is observed worldwide that the working-age population devote to the retired's social security in the PAYG program (Cremer, Gahvari, and Pestieau 2017). In addition, it is common in modern societies for the young to pay the medical care and climate change mitigation bills for the elderly, who tend to be more vulnerable to diseases and environmental hazards (Phaneuf and Requate 2017, Pestieau and Lefebvre 2018).

Consider an intergenerational contract, under which young adults devote a fraction $\tau \in (0, 1)$ of their incomes to provide their parents with public consumption. Such a contract can be rationalized even when each generation is selfish, ignoring the welfare of other generations (Rangel 2003). If members of one generation do not fulfill their contractual obligation, they will be unable to enjoy public consumption provided by their children when they become old, thereby suffering an infinitely negative utility (equation (2.1)). Ehrlich and Lui (1991) hold that an implicit social contract, in which a parent may invest in her children's education in exchange for their old-age support, can be sustained by intergenerational reputation.

In period t, the government and generation t engage in a Stackelberg game similar to that in Section 3. To maximize generation t's social welfare, the government moves first to provide public education and public consumption by collecting (net) taxes. After observing fiscal policies, each i-type member of generation t, chooses her private consumption during young and old adulthood (c_{yt}^i, c_{ot}^i) where $i \in \{s, u\}$.

Each i-type member of generation t chooses the private consumption levels to maximize her utility (2.1) subject to her budget constraint, which obtains

$$c_{yt}^s = \frac{(1 - \tau)w - \delta n - T_t^s}{1 + \alpha}, \qquad c_{ot}^s = \frac{\alpha(1 + r)[(1 - \tau)w - \delta n - T_t^s]}{1 + \alpha},$$
$$c_{yt}^u = \frac{1 - \tau - \delta n - T_t^u}{1 + \alpha}, \qquad c_{ot}^u = \frac{\alpha(1 + r)(1 - \tau - \delta n - T_t^u)}{1 + \alpha}. \qquad (4.1)$$

The government's balanced budget requires that the government spending on children's education and the elderly's public consumption are all born by young adults:

$$N_{t+1}E_t + N_{t-1}G_{t-1} = N_t\theta_t T_t^s + N_t(1 - \theta_t)T_t^u + \tau[w\theta_t N_t + (1 - \theta_t)N_t], \qquad (4.2)$$

where the intergenerational contract requires that[5]

$$N_{t-1}G_{t-1} = \tau[w\theta_t N_t + (1 - \theta_t)N_t]. \qquad (4.3)$$

---

[5]To circumvent technical difficulties, we assume that the intergenerational transfer is earmarked for public consumption of the elderly, which refers to a special PAYG system. With τ being a parameter, the economy may not achieve the social optimum. We discuss the optimal choice of τ at the end of this section.

Combining (4.2) and (4.3) shows that the government collects (T_t^s, T_t^u) from skilled and unskilled workers to finance the public education of their children, namely,

$$nE_t = \theta_t T_t^s + (1 - \theta_t) T_t^u. \tag{4.4}$$

The government maximizes the social welfare, V_t, subject to the budget constraint (4.2). Set up the Lagrangian as

$$
\begin{aligned}
\mathcal{L}_t = {} & \Delta + (1 + \alpha)\theta_t \ln[(1 - \tau)w - \delta n - T_t^s] + (1 + \alpha)(1 - \theta_t) \ln(1 - \tau - \delta n - T_t^u) \\
& + \beta[1 + (\rho - 1)\theta_t] \ln(1 + \sigma E_t) + \phi \ln\left\{n\tau[1 + (w - 1)\theta_{t+1}]\right\} \\
& - \mu_t \left[nE_t - \theta_t T_t^s - (1 - \theta_t)T_t^u\right],
\end{aligned}
\tag{4.5}
$$

with μ_t being the Lagrangian multiplier. The first order conditions with respect to E_t, T_t^s, T_t^u, and μ_t obtain the interior solutions as

$$
\begin{aligned}
\frac{\partial \mathcal{L}_t}{\partial E_t} &= \frac{\beta\sigma[1 + (\rho - 1)\theta_t]}{1 + \sigma E_t} + \frac{\sigma\phi(w - 1)[1 + (\rho - 1)\theta_t]}{(1 + \sigma E_t)[1 + (w - 1)\theta_{t+1}]} - n\mu_t = 0 \\
&\Rightarrow \mu_t = \frac{\sigma}{n} \cdot \frac{1 + (\rho - 1)\theta_t}{1 + \sigma E_t}\left[\beta + \frac{\phi(w - 1)}{1 + (w - 1)[1 + (\rho - 1)\theta_t]\ln(1 + \sigma E_t)}\right], \tag{4.6}
\end{aligned}
$$

$$
\frac{\partial \mathcal{L}_t}{\partial T_t^s} = -\frac{(1 + \alpha)\theta_t}{(1 - \tau)w - T_t^s - \delta n} + \mu_t\theta_t = 0 \quad \Rightarrow \quad (1 - \tau)w - \delta n - T_t^s = \frac{1 + \alpha}{\mu_t}, \tag{4.7}
$$

$$
\frac{\partial \mathcal{L}_t}{\partial T_t^u} = -\frac{(1 + \alpha)(1 - \theta_t)}{1 - \tau - T_t^u - \delta n} + \mu_t(1 - \theta_t) = 0 \quad \Rightarrow \quad 1 - \tau - \delta n - T_t^u = \frac{1 + \alpha}{\mu_t}, \tag{4.8}
$$

$$
\frac{\partial \mathcal{L}_t}{\partial \mu_t} = nE_t - \theta_t T_t^s - (1 - \theta_t)T_t^u = 0. \tag{4.9}
$$

Inserting (4.6) – (4.8) into (4.9) and rearranging characterizes E_t in the following function:

$$
\frac{(1 + \alpha)(1/\sigma + E_t)}{\frac{(1-\tau)[1+(w-1)\theta_t]}{n} - \delta - E_t} - \frac{\phi}{\frac{1}{(w-1)[1+(\rho-1)\theta_t]} + \ln(1 + \sigma E_t)} = \beta[1 + (\rho - 1)\theta_t]. \tag{4.10}
$$

The next proposition, which examines the traits of E_t, follows directly from (4.10):

Proposition 7 *The optimal educational expenditure per child in period t*

(a) increases with the proportion of skilled workers in generation t ($\frac{\partial E_t}{\partial \theta_t} > 0$);

(b) increases with the learning efficiency differential ($\frac{\partial E_t}{\partial \rho} > 0$);

(c) increases with the wage differential ($\frac{\partial E_t}{\partial w} > 0$);

(d) decreases with the fertility rate ($\frac{\partial E_t}{\partial n} < 0$).

Proof. See Appendix. ∎

Proposition 7 derives some comparative statics that are qualitatively similar to those in Propositions 1(a) and 3. Put it another way, the introduction of the intergenerational contract has little impact on our results of educational expenditure. It is noteworthy that in this section children are treated as not only a consumption good but also an investment good. While the quality of children is chosen by the government rather than the individual parent, the government does represent the interest of all parents as a whole. Becker-Lewis theorem continues to hold that an increase in the number of children lowers the quality of each child.

We proceed to investigate the properties of G_t in the proposition as below:

Proposition 8 *The optimal expenditure on public consumption per member of generation t*

(a) increases with the proportion of skilled workers in generation t ($\frac{\partial G_t}{\partial \theta_t} > 0$);

(b) increases with the learning efficiency differential ($\frac{\partial G_t}{\partial \rho} > 0$);

(c) increases with the wage differential ($\frac{\partial G_t}{\partial w} > 0$);

(d) increases with the fertility rate ($\frac{\partial G_t}{\partial n} > 0$).

Proof. See Appendix. ∎

Unlike Proposition 1(b), Proposition 8(a) presents a positive nexus between θ_t and G_t. Note that in section 3 one generation financially support their own public consumption, so they face a tradeoff between public education and public consumption. However, if public goods are supplied by the next generation – as in this section – such a tradeoff tends to fade. With economic development (larger θ_t), future generation is expected to consist of a larger proportion of skilled workers and a smaller proportion of unskilled workers. It is conceivable that as generation $t + 1$ earn a higher average income, generation t will receive a larger intergenerational transfer payment from their children, which they can spend on the consumption of public goods.

Propositions 8(b) and (c) have a straightforward intuition: a larger ρ or w heightens the income level of the next generation, thereby increasing the public consumption of the current generation. Also, the growing income inequality (larger w) increases the need for the government's social

security function. Proposition 8(d) shows that public consumption level is positively related to the fertility rate. This indicates that a high population growth rate can be an important reason for the introduction of this PAYG intergenerational contract.

We next turn to the steady state, as defined in Definition 1. For any given τ, the next proposition pins down the proportion of skilled workers in the working-age population in the steady state:

Proposition 9 *The steady-state proportion of skilled workers is characterized by*

$$e^{\frac{\theta}{1+(\rho-1)\theta}}\left[1+\frac{(1+\alpha)[1+\theta(w-1)]}{[1+(\rho-1)\theta][\beta+(\beta\theta+\phi)(w-1)]}\right]-\frac{\sigma[1+(w-1)\theta](1-\tau)}{n}=1-\sigma\delta. \quad (4.11)$$

Proof. See Appendix. ■

Proposition 9 demonstrates that in the steady state, θ is a function of nine parameters, namely, $(\alpha,\beta,\phi,\delta,\rho,\sigma,w,n,\tau)$. Comparing (4.11) with (3.13) suggests that the steady-state proportion of skilled workers in the working-age population under an intergenerational contract differs from that without such a contract. According to equation (4.11), the indicator of the intergenerational contract, τ, is a determinant of θ, implying that the value of τ tends to affect the long-term economic development; however, their relationship remains ambiguous.

An economy may not always treat the fraction of workers' income being transferred to parents as exogenous, but it can allow the people to make the optimal decision. We proceed to discuss the optimal intergenerational social contract, which maximizes the steady-state social welfare. In the steady state, equation (4.3) can be rewritten as

$$G = n\tau[1+(w-1)\theta]. \quad (4.12)$$

By (3.12), (4.6), and (4.12), we can rewrite V in (2.2) as a function of τ:

$$\begin{aligned}
V(\tau) &= \alpha\ln[\alpha(1+r)]-(1+\alpha)\ln\left\{\frac{\sigma}{n}\cdot\frac{1+(\rho-1)\theta}{1+\sigma E}\left[\beta+\frac{\phi(w-1)}{1+\theta(w-1)}\right]\right\}\\
&\quad +\beta[(\rho-1)\theta+1]\ln(1+\sigma E)+\phi\ln\left\{n\tau[1+(w-1)\theta]\right\}\\
&= \Pi+\phi\ln\tau+\beta\theta+\frac{(1+\alpha)\theta}{1+(\rho-1)\theta}+(1+\alpha+\phi)\ln[1+(w-1)\theta]\\
&\quad -(1+\alpha)\ln[\beta+(\phi+\beta\theta)(w-1)]-(1+\alpha)\ln[1+(\rho-1)\theta]. \quad (4.13)
\end{aligned}$$

17

where $\Pi := \alpha \ln[\alpha(1+r)] + (1 + \alpha + \phi) \ln n - (1 + \alpha) \ln \sigma$ is a constant, while $\theta = \theta(\tau)$ is solved in Proposition 9. By denoting $\tau^* = \arg\max\{V\}$, we establish the next proposition to show the existence of τ^*:

Proposition 10 *The optimal intergenerational contract τ^* exists.*

Proof. See Appendix. ■

The intergenerational social contract essentially reflects a trade between different generations, which can be sustained by a self-fulfilling equilibrium. If people contribute more to their parents for their public consumption, they expect to enjoy more public consumption contributed by their children when they become old; but they also have to cut their private consumption and expenses on educating children. The optimal intergenerational contract, τ^*, which falls in the range $[0, 1]$ and maximizes the *ex ante* utility of a representative individual in the long run, is characterized by equations (4.11) and (4.13).

5   Simulation: A Comparative Analysis

We have thus far provided a theoretical analysis of the optimal allocation of public expenditures among different functions with and without an intergenerational social contract. A natural question arises as to which regime leads to a higher social welfare? The answer to this question rests on the parameter configuration. In this section, we conduct a quantitative exercise to simulate the changes in educational expenditure and in public consumption over time under the two regimes, and illustrate the welfare impact of the introduction of the intergenerational contract in both transitional dynamics and steady state.

To implement the simulation, we assign values to seven of our model's parameters, namely, $(\alpha, \beta, \phi, \delta, \rho, \sigma, r)$, as listed in Table 2. In particular, we set $\delta = 1$, which means that an unskilled worker spends 5% of her wage to bear and raise a child. The learning efficiency differential $\rho = 5$ indicates that a child with a skilled parent is five times more likely to be skilled in the future than a child with an unskilled parent. The (intertemporal) interest rate is set to be one, which can be translated into an annual compound interest rate of 3% over 24 years or an annual compound interest rate of 2.5% over 28 years. Furthermore, let the economy begin with a relatively small

Table 2. Parameter Values in Simulation

Parameter	Interpretation	Value
α	Preference over private consumption in old adulthood	0.1
β	Preference over child quality	100
ϕ	Preference over public consumption	0.3
δ	Cost of rearing one child	0.05
ρ	Learning efficiency differential	5
σ	Efficiency of educational expenditure	0.1
r	Interest rate	1
θ_0	Proportion of skilled workers in the initial period	0.01

proportion of skilled workers so that we can better observe the economic development. We thus set the initial value of θ to be 1 percent (i.e., $\theta_0 = 0.01$).

There are two additional key parameters: the fertility rate n and the wage differential w. In our numerical exercises, we consider two sets of values: (1) each individual bears 1.4 children on average, and a skilled worker's wage is three times as high as an unskilled worker's wage ($n = 1.4$, $w = 3$), and (2) each individual bears 1.6 children on average, and a skilled worker's wage is four times as high as an unskilled worker's wage ($n = 1.6$, $w = 4$). Given the above information, we compute the optimal intergenerational contract as $\tau^* \approx 0.02$.

Figure 1 illustrates the first set of simulation results given a fertility rate of $n = 1.4$ and a wage differential of $w = 3$. The dark bars represent the absence of an intergenerational contract (section 3) and the gray bars represent the presence of such a contract (section 4). There are three plots, each of which displays the evolution of educational expenditure per child (E_t), public consumption per member of generation t (G_t), and social welfare (V_t).
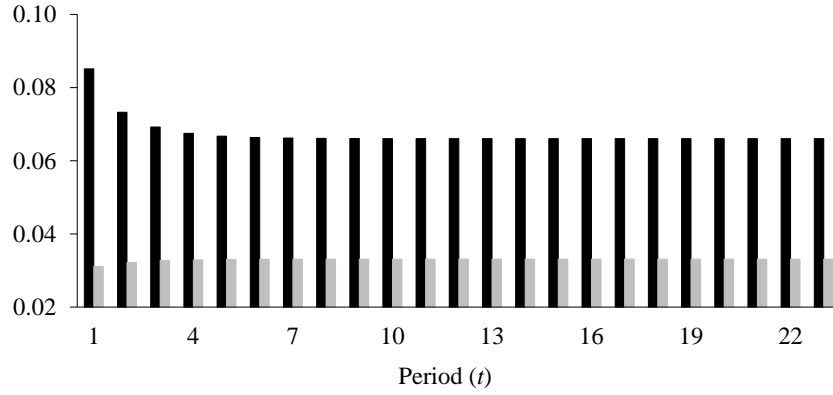
Plot (a) charts that both dark and gray bars exhibit an increasing trend in height. As the economy begins with 1 percent of skilled young workers, the growing educational expenditure raises the proportion of skilled workers over time (equation (2.6)), giving rise to continuous economic development. Gray bars are taller than dark bars, signifying that an intergenerational contract stimulates educational spending. In the long run, the height of gray bars remains at $E = 0.692$, with a

Figure 1: Dynamics of Public Expenditure and Social Welfare ($n = 1.4$ and $w = 3$)
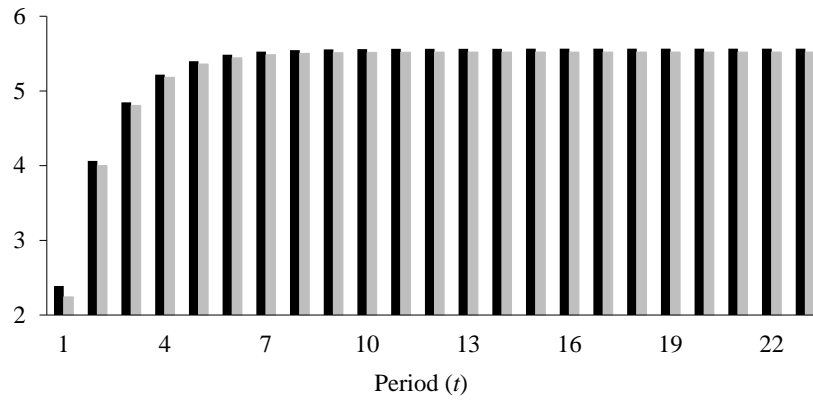
(a) Expenditure on Public Education (E_t)

(b) Expenditure on Public Consumption (G_t)

(c) Social Welfare (V_t)

■ No Intergenerational Contract   ▧ Intergenerational Contract

proportion $\theta = 9.14\%$ of the working-age population being skilled workers. In contrast, dark bars show that the educational spending is kept at $E = 0.682$ in the steady state, which corresponds to a proportion $\theta = 8.97\%$ of skilled workers.

Although the introduction of a pay-as-you go intergenerational contract facilitates human capital accumulation, it may undermine the provision of public goods. In Plot (b), dark and gray bars display different trends. The intergenerational contract allows every generation to consume more public goods than their parents did until the steady-state level of public consumption $G = 0.033$ is achieved. In contrast, the absence of intergenerational transfers leads to regressive public consumption, largely due to the crowding-out effect of the expanding educational expenditure. Specifically, public consumption per capita decreases from 0.085 in period 1 to 0.066 in period 23. This outcome echoes the prediction of Proposition 1(b).
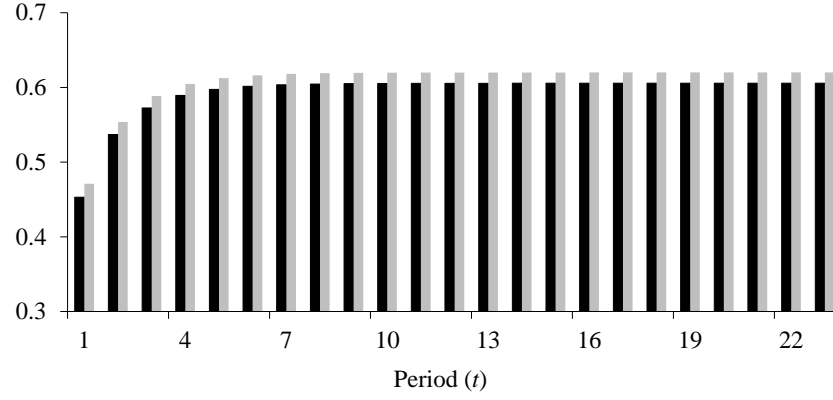
Plot (c) of Figure 1 compares the evolution of social welfare with and without intergenerational contract. Both dark and gray bars display a rising trend in height, indicating that the future generation will enjoy a higher social welfare than the current generation. Nevertheless, dark bars are always taller than gray bars, although their gaps become narrower when the economy approaches to the steady state. Interpreted literally, the absence of intergenerational contract tends to be welfare-improving.

We redo the simulation in another configuration in which the fertility rate and the wage differential are set to higher values. Figure 2 charts the set of simulation results contrasting E_t, G_t, and V_t in the two regimes given the fertility rate $n = 1.6$ and the wage differential $w = 4$. In plot (a), the educational expenditure increases over time, which parallels the growth in the proportion of skilled workers until a steady state is achieved. Plot (b) illustrates that the implementation of the optimal intergenerational contract results in a lower steady-state level of public consumption. These two outcomes are almost the same as in Figure 1(a) and 1(b).
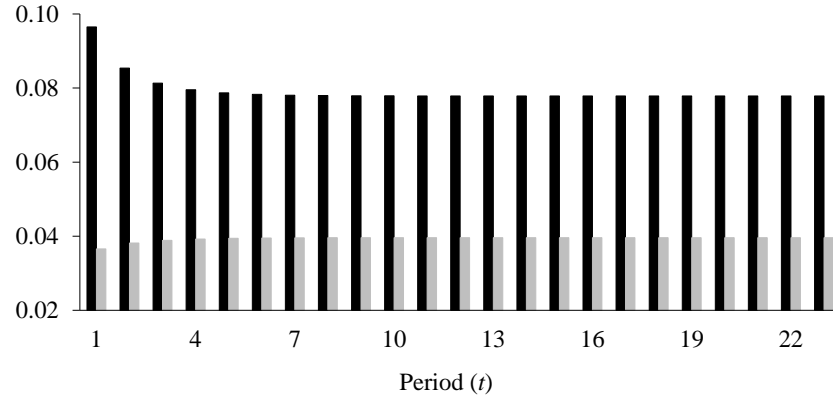
In plot (c) of Figure 2, the heights of both dark and gray bars increase over time and approach to their respective steady-state levels. To be concrete, social welfare increases from 1.73 in period 1 to 4.51 in period 23 without intergenerational transfers, yet it increases from 1.62 in period 1 toward 4.53 otherwise. Dark bars are generally taller than gray bars in the first two periods and fall shorter from the third period onward. In the long run, social welfare is higher when the optimal intergenerational contract is implemented. We can see that a relatively high population growth rate

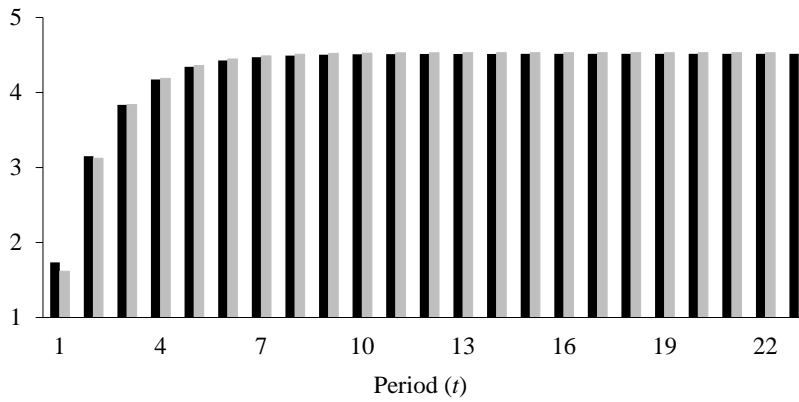Figure 2: Dynamics of Public Expenditure and Social Welfare ($n = 1.6$ and $w = 4$)

(a) Expenditure on Public Education (E_t)

(b) Expenditure on Public Consumption (G_t)

(c) Social Welfare (V_t)

■ No Intergenerational Contract   ▨ Intergenerational Contract

allows each young individual to bear a smaller share of cost in supporting public consumption of the old, which may lead to a dynamic Pareto improvement for all generations.

In most literature, $1 + r < n$ is a simple criterion assessing the welfare-improving effect of a PAYG program. Our simulation example suggests that even in the case $1 + r > n$, an intergenerational contract may improve social welfare. This is largely because in our model with a PAYG program (section 4), children are treated as an investment good: parents financially support public education for their children, who in turn finance parent's public consumption in future. Given the assumption that children provide their parents with public consumption, parents are inclined to spend more on children's education, which creates mutual benefits to both generations.

6 Robustness of the Model

In this section, we aim to examine the robustness of our model by relaxing some simplifying assumptions and discussing potential extensions. Specifically, we explore a more profound relationship between a parent's skill and her child's skill in section 6.1 and allow for a time-varying relative wage of skilled workers in section 6.2. We then consider the distortionary effect of tax collection on labor supply in section 6.3 and endogenous fertility in section 6.4.

6.1 Family Background and Children's Human Capital

We assume the parameter $\rho > 1$ as in equation (2.3), meaning that children with skilled parents have an advantage in accumulating human capital and becoming skilled than those with unskilled parents. The importance of parental human capital in shaping a child's human capital has been consistently documented in the literature (See Hanushek (1996) for a survey). In addition to parental skill and school education, some other elements, such as parents' time and effort in educating their children, also matter to children's human capital. Many empirical studies find a positive influence of home education on children's academic achievements (Behrman, Foster, Rosenzweig, and Vashishtha 1999, Ramey and Ramey 2010). A person's cognitive skill is largely developed in the early environment created and maintained by her parents (Heckman, Pinto, and Savelyev 2013, Almond, Currie, and Duque 2018). There is also supportive evidence that home environment affects children's learning efficiency at school (Vernon 1979).

Based on the above analysis, we now relax our assumption by formulating a more generalized expression of a child's quality in period t as

$$q_t^i = q(\chi_{t-1}^i, E_t, h_t^i), \tag{6.1}$$

where χ_{t-1}^i is an indicator of parental human capital, h_t^i denotes the time that an i-type member of generation t spends in educating her child (i.e., home education). Parents' time in educating their children entails a cost: an increase in h_t^i decreases an i-type parent's working time and hence her earnings.

Equation (6.1) or a similar formulation is often used in earlier theoretical literature (e.g., Benabou 1996, De la Croix and Doepke 2003, 2004). This literature suggests that there are often two opposite effects shaping the choice of h_t^i. On one hand, a parent with greater human capital is more efficient in undertaking home education so that she would devote more to educating children. On the other hand, a parent with greater human capital receives a higher wage, which means that she faces a larger opportunity cost of educating children. In equilibrium, these two effects often cancel out each other, leaving all parents spend the same amount of time in educating their children (i.e., h_t^i turns out to be constant across population and thus independent of i).

Equation (2.3) shows that there are two different levels of parental human capital only (i.e., the skilled and the unskilled), which can be regarded as a reduced form of (6.1). We may simplify (6.1) as follows:

$$q_t^i = p(\chi_{t-1}^i, h_t^i) \ln(1 + \sigma E_t). \tag{6.2}$$

Cremer, De Donder, and Pestieau (2010) develop a two-skill setting where skilled parents are more likely to have skilled children than unskilled parents in the presence of elitism of public education. Following their study, we normalize the learning efficiency of children with unskilled parents to be one and set the learning efficiency of children with skilled parents to be greater than one, i.e., $p(\chi_{t-1}^i, h_t^i) = 1$ if $i = u$ and $p(\chi_{t-1}^i, h_t^i) = \rho > 1$ if $i = s$. From the above reasoning, we can see that our modeling strategy, which captures the dependence of a child's quality on parental human capital, is a reasonable simplified form of a more generalized formulation.

24

6.2 Time-Varying Wage Differential

In our simple model, we consider a small open economy where the relative wage of skilled to unskilled workers is exogenous and constant over time. Yet a more realistic and generalized setting may be that the relative wage is endogenous and time-dependent. This section studies the case that the wage differential of one generation is governed by the proportion of skilled workers:

$$w_t = w(\theta_t). \tag{6.3}$$

The sign of $w'(\theta_t)$ may be ambiguous. On one hand, the law of diminishing marginal returns implies that holding other things fixed, the increased number of skilled workers results in a lower marginal product of skilled labor and hence a lower wage rate, namely, $w'(\theta_t) < 0$. When the proportion of skilled workers is high, the welfare-maximizing government tends to curb its adverse impact on the wage rate of skilled workers by limiting the increase in skilled labor force. Accordingly, the government needs to allocate resources from public education toward other expenditure categories.

On the other hand, the increased number of skilled workers may create the agglomeration of human capital, which in turn facilitates their learning from each other and improves labor productivity. As such, the more the skilled workers, the higher their wage rate, i.e., $w'(\theta_t) > 0$. This positive link is not divorced from empirical evidence. For example, Rosenthal and Strange (2008) find that the human capital concentration benefits wage growth. Karlsson and Gråsjö (2014) demonstrate that such a beneficial effect is prominent in high-tech knowledge-intensive firms. Earlier theoretical studies have analyzed the importance of human capital externalites in the determination of wage rate. Lucas (1988) and Becker, Murphy, and Tamura (1990) introduce the average human capital into the production function showing that labor productivity increases with the average labor skill. Fan and Stark (2008) also specify the positive association between skilled workers' wage rate and their share in labor force. Given that positive externalities are sufficiently large, as in a knowledge-based economy, the government may find it desirable to contribute more resources to public education even if skilled workers account for a large proportion of labor force.

6.3 Costs of Taxation and Redistribution

Our simple model abstracts from informational asymmetry and administrative costs of taxation and government expenditures. Under this assumption, we develop Proposition 5 that all individuals have the same after-tax income level, namely, $w - T^s = 1 - T^u$. However, this result may not hold if tax collection entails a cost. Specifically, the government's design of the income tax rate and/or transfer payment may induce a reduction of labor supply or tax evasion behaviors.

Since the seminal contribution of Mirrlees (1971), there has been vast literature examining how income taxation affects people's labor supply and social welfare given asymmetric information. While a benevolent government has an incentive to mitigate wage inequality by levying an income tax on high earners, it is also aware that an overly high tax discourages labor supply and dampens productivity (e.g., Cremer, Pestieau, and Rochet 2001). Therefore, the government needs to strike a balance between redistributing income from the rich to the poor and reducing the disincentive effect of taxes on labor supply.

Moreover, there are often substantial administrative costs associated with taxation and redistribution. For example, a high tax rate may trigger strategic tax evasions in an underground economy (e.g., Allingham and Sandmo 1972, Slemrod 2007). It is often a challenging task facing the government to find the desirable formula to ensure that taxation brings in enough revenues to pay for a wide range of public services without demoralizing taxpayers.

Considering the potential fall in labor supply and tax evasion, the government tends to lower income tax.[6] As such, a skilled worker may get a higher after-tax income than an unskilled worker, and the income gap is an increasing function of the cost associated with taxation. Because informational asymmetry and administrative costs tend to decrease with economic development, a more advanced economy may place a greater emphasis on taxation and redistribution.

6.4 Endogenous Fertility Choices

Our model considers that all parents have the same number of children for simplification. Much received literature analyzes endogenous fertility. For example, Becker and Lewis (1973) posit that

---

[6]Auerbach and Hines (2002) and Piketty and Saez (2013) provide surveys on the optimal taxation evaluating these tradeoffs.

parents, who obtain utility from both quantity and quality of their children, may substitute quality for quantity under certain conditions. Willis (1973) extends Becker-Lewis model showing that the time and pecuniary constraints influence parents' choices of the quantity and quality of their children. De la Croix and Doepke (2003, 2004) address the negative link between a person's income and fertility, suggesting that this fertility-differential effect accounts for most of the empirical relationship between inequality and development. Other related literature on endogenous fertility includes Galor and Weil (2000), Fan and Zhang (2013), and Varvarigos and Zakaria (2013).

We proceed to discuss how the negative income-fertility relationship influences the main result of our model. When the skilled have fewer children than the unskilled, it is welfare-improving to reallocate educational resources from the rich to the poor by the law of diminishing returns. Besides, with an increased educational expenditure in the current period, there will be a greater proportion of skilled workers in the next period, leading to a lower average fertility rate. In that case, future generations will receive a higher level of educational expenditure per person, which constitutes a virtuous cycle. Overall, allowing for endogenous and differential fertility may reinforce the importance of public education and the income redistribution.

7 Conclusion

From the perspective of public finance, tax revenue and government spending are two sides of the same coin. This paper adds to the literature on the optimal allocation of government spending by incorporating individual heterogeneity and intergenerational mobility. In our model, a government plans its expenditures according to the objectives of social protection, public education, and consumption of public goods and services, each addressing a different aspect of social welfare. Total public revenues, which come from income taxes, impose a constraint that necessitates trade-offs between these competing objectives, thereby affecting the composition of government spending.

Our model considers individual heterogeneity in terms of skill in a dynamic setting in which individuals differ not only in earning capacity but also in the intra-family influence on children's learning efficiency. We first analyze the case that each generation maximizes their own expected utility without making direct wealth transfers to their parents. Our analysis suggests that in the steady state where the proportion of skilled individuals remains constant, the educational expen-

diture increases with the learning efficiency of skilled workers' children and the wage differential but decreases with the fertility rate. Also, under some configurations, expenditure on public consumption is inversely related to the learning efficiency of skilled workers's children but positively related to the wage differential and the fertility rate.

We then investigate the optimal allocation of government expenditures with intergenerational wealth transfers. If there exists a PAYG program allowing that the young generation devote a certain proportion of their income to support the public consumption of the old generation, our model delivers some different implications. By examining the transitional dynamics, we show that educational expenditure and public consumption tend to increase with economic development. A higher fertility rate leads to a cut in educational expenditure yet an increase in public consumption. We also characterize the intergenerational contract that maximizes the steady-state social welfare. Finally, we conduct a quantitative simulation comparing the welfare consequences of the regime without an intergenerational contract and that of the regime with the optimal intergenerational contract. Our simulation exercise illustrates that the introduction of optimal intergenerational contract may improve social welfare.

Appendix: Mathematical Proofs

Proof of Proposition 1

(a) Differentiating E_t in (3.7) with respect to θ_t derives

$$\frac{\partial E_t}{\partial \theta_t} = \frac{(w-1)\left(1 + \frac{1+\alpha+\phi}{\beta[1+(\rho-1)\theta_t]}\right) + \frac{(\rho-1)(1+\alpha+\phi)[1+(w-1)\theta_t+(1/\sigma-\delta)n]}{\beta[1+(\rho-1)\theta_t]^2}}{n\left(1 + \frac{1+\alpha+\phi}{\beta[1+(\rho-1)\theta_t]}\right)^2} > 0. \qquad \text{(A.1)}$$

(b) Differentiating G_t in (3.8) with respect to θ_t obtains

$$\frac{\partial G_t}{\partial \theta_t} = \frac{\phi(1+r)[(w-1)(1+\alpha+\beta+\phi) - \beta(\rho-1)(1-\delta n + n/\sigma)]}{[1+\alpha+\beta+\phi+\beta(\rho-1)\theta_t]^2}, \qquad \text{(A.2)}$$

which is negative if and only if $(w-1)(1+\alpha+\beta+\phi) < \beta(\rho-1)(1-\delta n + n/\sigma)$, namely, condition (3.11) holds.

Proof of Proposition 2

Combining (3.7) and (3.12) obtains that in the steady state,

$$\frac{e^{\frac{\theta}{1+(\rho-1)\theta}} - 1}{\sigma} = \frac{\beta[1+(\rho-1)\theta][1+(w-1)\theta-\delta n]/n - (1+\alpha+\phi)/\sigma}{1+\alpha+\beta+\phi+\beta(\rho-1)\theta}$$

$$\Rightarrow \frac{\theta}{1+(\rho-1)\theta} = \ln\left\{\frac{\beta[1+(\rho-1)\theta][\sigma/n+\sigma(w-1)\theta/n-\delta\sigma] - (1+\alpha+\phi)}{1+\alpha+\beta+\phi+\beta(\rho-1)\theta} + 1\right\}, \text{ (A.3)}$$
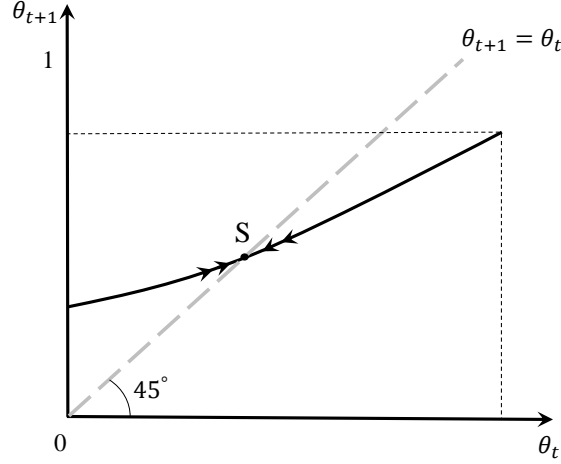
which is equivalent to (3.13). We can infer from (2.6) that

$$\text{For } \theta_t = 1, \quad \theta_{t+1} = \frac{q^s n \theta_t N_t + 0}{N_{t+1}} = q^s < 1 = \theta_t; \qquad \text{(A.4)}$$

$$\text{For } \theta_t = 0, \quad \theta_{t+1} = \frac{0 + q_t^u n N_t}{N_{t+1}} = q_t^u > 0 = \theta_t. \qquad \text{(A.5)}$$

The above two equations demonstrate that the proportion of skilled workers decreases over time if it starts from 100% but increases over time if it starts from zero. Given that θ_{t+1} is continuous in θ_t over the closed interval $[0,1]$, the intermediate value theorem suggests that there exists $\theta_t =$

Figure 3: The Locus of the Proportion of Skilled Workers

$\theta_{t+1} = \theta \in (0,1)$, which determines the (nontrival) steady state. Differentiating θ_{t+1} in equation (2.6) with respect to θ_t and using Proposition 1(a) yields

$$\frac{d\theta_{t+1}}{d\theta_t} = (\rho - 1)\ln(1 + \sigma E_t) + \frac{[1 + (\rho - 1)\theta_t]\sigma E_t}{1 + \sigma E_t}\frac{dE_t}{d\theta_t} > 0. \qquad (A.6)$$

Figure 3 presents a graphic illustration of the locus of θ_t in the space (θ_t, θ_{t+1}), where points (θ_t, θ_{t+1}) for all t are interconnected by a dark curve. This curve passes through the straight dashed line $\theta_{t+1} = \theta_t$ (the 45 degree line) from above. This crossing point corresponds to the steady state (as represented by S), at which the locus has a slope of

$$0 < \frac{d\theta_{t+1}}{d\theta_t} < 1. \qquad (A.7)$$

Therefore, the point S is a stable equilibrium, to which the nearby solutions converge asymptotically. We draw arrows to describe the direction of motion near S.

Proof of Proposition 3

By (3.13), we introduce the following notation:

$$Z(\theta; \rho, w, n) := [1 + (\rho - 1)\theta]\ln\left\{\frac{1 - \delta\sigma + \sigma/n + (w - 1)\theta\sigma/n}{1 + \frac{1+\alpha+\phi}{\beta[1+(\rho-1)\theta]}}\right\} - \theta = 0. \qquad (A.8)$$

30

It is straightforward to derive from (A.8) that

$$\frac{\partial Z}{\partial \rho} > 0, \qquad \frac{\partial Z}{\partial n} < 0, \qquad \frac{\partial Z}{\partial w} > 0. \tag{A.9}$$

By inserting (A.8), we rewrite (2.6) to obtain $Z(\theta_t) = \theta_{t+1} - \theta_t$. Differentiating Z with respect to θ_t and then using (A.7) yields

$$\frac{dZ(\theta_t)}{d\theta_t} = \frac{d\theta_{t+1}}{d\theta_t} - 1 < 0. \tag{A.10}$$

Totally differentiating (A.8) and then using (A.9) and (A.10) obtains

$$\frac{\partial Z}{\partial \theta}d\theta + \frac{\partial Z}{\partial \rho}d\rho = 0 \quad \Rightarrow \quad \frac{d\theta}{d\rho} = -\frac{\frac{\partial Z}{\partial \rho}}{\frac{\partial Z}{\partial \theta}} > 0, \tag{A.11}$$

$$\frac{\partial Z}{\partial \theta}d\theta + \frac{\partial Z}{\partial n}dn = 0 \quad \Rightarrow \quad \frac{d\theta}{dn} = -\frac{\frac{\partial Z}{\partial n}}{\frac{\partial Z}{\partial \theta}} < 0, \tag{A.12}$$

$$\frac{\partial Z}{\partial \theta}d\theta + \frac{\partial Z}{\partial w}dw = 0 \quad \Rightarrow \quad \frac{d\theta}{dw} = -\frac{\frac{\partial Z}{\partial w}}{\frac{\partial Z}{\partial \theta}} > 0. \tag{A.13}$$

Differentiating (3.12) with respect to θ shows that E and θ are positively related:

$$\frac{\partial E}{\partial \theta} = \frac{1}{\sigma} \cdot e^{\frac{\theta}{1+(\rho-1)\theta}} \cdot \frac{1 + (\rho-1)\theta - \theta(\rho-1)}{[1+(\rho-1)\theta]^2} = \frac{1+\sigma E}{\sigma[(\rho-1)\theta+1]^2} > 0. \tag{A.14}$$

Differentiating E in (3.7) with respect to ρ and using (3.8) yields

$$\begin{aligned}
\frac{dE}{d\rho} &= \frac{\beta\theta(1+\alpha+\phi)[1+(w-1)\theta+(1/\sigma-\delta)n]}{n[1+\alpha+\beta+\phi+\beta(\rho-1)\theta]^2} + \frac{\partial E}{\partial \theta}\frac{\partial \theta}{\partial \rho} \\
&= \frac{\beta\theta(1+\alpha+\phi)G}{n\phi(1+r)[1+\alpha+\beta+\phi+\beta(\rho-1)\theta]} + \frac{\partial E}{\partial \theta}\frac{\partial \theta}{\partial \rho} > 0, \tag{A.15}
\end{aligned}$$

because the first term is positive in equilibrium, and the second term is positive by equations (A.11) and (A.14).

Differentiating E in (3.7) with respect to n yields

$$\frac{dE}{dn} = -\frac{\beta[1+(\rho-1)\theta][1+(w-1)\theta]}{[1+\alpha+\beta+\phi+\beta(\rho-1)\theta]n^2} + \frac{\partial E}{\partial \theta}\frac{\partial \theta}{\partial n} < 0, \tag{A.16}$$

because the first term is negative, and the second term is also negative by (A.12) and (A.14).

Differentiating E in (3.7) with respect to w obtains

$$\frac{dE}{dw} = \frac{\theta}{n\left(1 + \frac{1+\alpha+\phi}{\beta[1+(\rho-1)\theta]}\right)} + \frac{\partial E}{\partial \theta}\frac{\partial \theta}{\partial w} > 0, \tag{A.17}$$

because the first term is positive, and the second term is positive by equations (A.13) and (A.14).

Proof of Proposition 4

In the steady state, the public consumption per old adult in (3.8) can be written as

$$G = \frac{\phi(1+r)[1+(w-1)\theta+(1/\sigma-\delta)n]}{1+\alpha+\beta+\phi+\beta(\rho-1)\theta}. \tag{A.18}$$

Differentiating G in (A.18) with respect to θ obtains

$$\frac{\partial G}{\partial \theta} = \frac{\phi(1+r)[(w-1)(1+\alpha+\beta+\phi)-\beta(\rho-1)(1-\delta n+n/\sigma)]}{[1+\alpha+\beta+\phi+\beta(\rho-1)\theta]^2}, \tag{A.19}$$

which is nonpositive if and only if condition (3.11) holds.

(a) Differentiating G in (A.18) with respect to ρ obtains

$$\frac{dG}{d\rho} = -\frac{\beta\phi(1+r)[1+(w-1)\theta+(1/\sigma-\delta)n]\theta}{[1+\alpha+\beta+\phi+\beta(\rho-1)\theta]^2} + \frac{\partial G}{\partial \theta}\frac{\partial \theta}{\partial \rho}, \tag{A.20}$$

where the first term is negative, and the second term is nonpositive by (A.11) and (A.19) if and only if condition (3.11) holds.

(b) Differentiating G in (A.18) with respect to n obtains

$$\frac{dG}{dn} = \frac{\phi(1+r)(1/\sigma-\delta)}{1+\alpha+\beta+\phi+\beta(\rho-1)\theta} + \frac{\partial G}{\partial \theta}\frac{\partial \theta}{\partial n}, \tag{A.21}$$

where the first term is positive if $\delta\sigma < 1$, and the second term is positive by (A.12) and (A.19) if and only if condition (3.11) holds.

(c) Differentiating G in (A.18) with respect to w obtains

$$\frac{dG}{dw} = \frac{\phi(1+r)\theta}{1+\alpha+\beta+\phi+\beta(\rho-1)\theta} + \frac{\partial G}{\partial \theta}\frac{\partial \theta}{\partial w}, \tag{A.22}$$

32

where the first term is always positive, and the second term is positive if and only if condition (3.11) is not satisfied.

Proof of Proposition 6

Differentiating T^s and T^u with respect to θ obtains

$$\frac{\partial T^s}{\partial \theta} = \frac{\partial T^u}{\partial \theta} = -\frac{(1+\alpha)[(w-1)(1+\alpha+\beta+\phi) - \beta(\rho-1)(1-\delta n + n/\sigma)]}{[1+\alpha+\beta+\phi+\beta(\rho-1)\theta_t]^2}, \quad (A.23)$$

which is positive if and only if (3.11) holds.

(a) Differentiating T^s and T^u with respect to ρ obtains

$$\frac{dT^s}{d\rho} = \frac{dT^u}{d\rho} = \frac{\beta\theta(1+\alpha)[1+(w-1)\theta+(1/\sigma-\delta)n]}{[1+\alpha+\phi+\beta(1+\rho\theta-\theta)]^2} + \frac{\partial T^s}{\partial\theta}\frac{\partial\theta}{\partial\rho}, \quad (A.24)$$

where the first term is positive, and the second term is positive if and only if condition (3.11) is satisfied.

(b) Differentiating T^s and T^u with respect to n yields

$$\frac{dT^s}{dn} = \frac{dT^u}{dn} = -\frac{(1+\alpha)(1/\sigma-\delta)}{\sigma[1+\alpha+\beta+\phi+\beta(\rho-1)\theta]} + \frac{\partial T^s}{\partial\theta}\frac{\partial\theta}{\partial n}, \quad (A.25)$$

where the first term is negative if and only if $\sigma\delta < 1$, and second term is negative by (A.12) and (A.23) if and only if (3.11) holds.

(c) Differentiating T^s with respect to w obtains

$$\frac{dT^s}{dw} = \frac{(1+\alpha)(1-\theta)+\beta+\phi+\beta(\rho-1)\theta}{1+\alpha+\beta+\phi+\beta(\rho-1)\theta} + \frac{\partial T^s}{\partial\theta}\frac{\partial\theta}{\partial w}, \quad (A.26)$$

where the first term is positive, and the second term is positive if and only if (3.11) holds.
Differentiating T^u with respect to w obtains

$$\frac{dT^u}{dw} = -\frac{(1+\alpha)\theta}{1+\alpha+\beta+\phi+\beta(\rho-1)\theta} + \frac{\partial T^u}{\partial\theta}\frac{\partial\theta}{\partial w}, \quad (A.27)$$

where the first term is negative, and the second term is negative if and only if (3.11) is not satisfied.

Proof of Proposition 7

(a) Rewrite (4.10) as $g(\theta_t, E_t) = 0$. Differentiating g with respect to E_t and θ_t obtains

$$
\frac{\partial g}{\partial E_t} = (1+\alpha)\left\{ \frac{\left[\frac{(1-\tau)[1+(w-1)\theta_t]}{n} - \delta - E_t\right] + (1/\sigma + E_t)}{\left[\frac{(1-\tau)[1+(w-1)\theta_t]}{n} - \delta - E_t\right]^2} \right\}
$$
$$
+ \frac{\phi}{\left[\frac{1}{(w-1)[1+(\rho-1)\theta_t]} + \ln(1+\sigma E_t)\right]^2} \frac{E_t}{1+\sigma E_t} > 0, \tag{A.28}
$$

$$
\frac{\partial g}{\partial \theta_t} = -\frac{(1+\alpha)(1/\sigma + E_t)}{\left[\frac{(1-\tau)[1+(w-1)\theta_t]}{n} - \delta - E_t\right]^2} \frac{(1-\tau)(w-1)}{n} - \beta(\rho - 1)
$$
$$
- \frac{\phi(\rho-1)(w-1)}{[1+(w-1)[1+(\rho-1)\theta_t]\ln(1+\sigma E_t)]^2} < 0. \tag{A.29}
$$

Totally differentiating g with respect to E_t and θ_t and rearranging yields

$$
\frac{\partial g}{\partial \theta_t}d\theta_t + \frac{\partial g}{\partial E_t}dE_t = 0 \quad \Leftrightarrow \quad \frac{dE_t}{d\theta_t} = -\frac{\partial g/\partial \theta_t}{\partial g/\partial E_t} > 0. \tag{A.30}
$$

(b) Rewrite (4.10) as $y(\rho, E_t) = 0$, where $\frac{\partial y}{\partial E_t} = \frac{\partial g}{\partial E_t} > 0$. Moreover, we have

$$
\frac{\partial y}{\partial \rho} = \frac{\phi}{\left\{\frac{1}{(w-1)[1+(\rho-1)\theta_t]} + \ln(1+\sigma E_t)\right\}^2} \frac{1}{w-1} \cdot \frac{-\theta_t}{[1+(\rho-1)\theta_t]^2} - \beta\theta_t
$$
$$
= -\left\{ \frac{\phi(w-1)\theta_t}{\{1+(w-1)[1+(\rho-1)\theta_t]\ln(1+\sigma E_t)\}^2} + \beta\theta_t \right\} < 0. \tag{A.31}
$$

Totally differentiating y with respect to ρ and E_t and rearranging yields

$$
\frac{\partial y}{\partial \rho}d\rho + \frac{\partial y}{\partial E_t}dE_t = 0 \quad \Leftrightarrow \quad \frac{dE_t}{d\rho} = -\frac{\partial y/\partial \rho}{\partial y/\partial E_t} > 0. \tag{A.32}
$$

(c) Rewrite (4.10) as $z(n, E_t) = 0$, where $\frac{\partial z}{\partial E_t} = \frac{\partial g}{\partial E_t} > 0$. Moreover, we have

$$
\frac{\partial z}{\partial n} = \left(\frac{\partial g}{\partial E_t}\right)^{-1} \frac{(1+\alpha)(1/\sigma + E_t)(1-\tau)[1+(w-1)\theta_t]}{\{(1-\tau)[1+(w-1)\theta_t] - (\delta + E_t)n\}^2} > 0. \tag{A.33}
$$

Totally differentiating z with respect to n and E_t yields

$$\frac{\partial z}{\partial n}dn + \frac{\partial z}{\partial E_t}dE_t = 0 \qquad \Leftrightarrow \qquad \frac{dE_t}{dn} = -\frac{\partial z/\partial n}{\partial z/\partial E_t} < 0. \tag{A.34}$$

(d) Rewrite (4.10) as $h(w, E_t) = 0$, where $\dfrac{\partial h}{\partial E_t} = \dfrac{\partial g}{\partial E_t} > 0$. Moreover, we have

$$\frac{\partial h}{\partial w} = \frac{-(1+\alpha)(1/\sigma + E_t)(1-\tau)\theta_t}{n\left[\frac{(1-\tau)[1+(w-1)\theta_t]}{n} - \delta - E_t\right]^2} - \frac{\phi[1+(\rho-1)\theta_t]}{\{1+(w-1)[1+(\rho-1)\theta_t]\ln(1+\sigma E)\}^2} < 0. \tag{A.35}$$

Totally differentiating h with respect to w and E_t yields

$$\frac{\partial h}{\partial w}dw + \frac{\partial h}{\partial E_t}dE_t = 0 \qquad \Leftrightarrow \qquad \frac{dE_t}{dw} = -\frac{\partial h/\partial w}{\partial h/\partial E_t} > 0. \tag{A.36}$$

Proof of Proposition 8

Plugging (2.6) into (4.3) obtains

$$G_t = n\tau\{1 + (w-1)[1+(\rho-1)\theta_t]\ln(1+\sigma E_t)\}. \tag{A.37}$$

Differentiating it with respect to θ_t, ρ, and w obtains

$$\begin{aligned}
\frac{dG_t}{d\theta_t} &= n\tau(w-1)\left[(\rho-1)\ln(1+\sigma E_t) + \frac{\sigma[1+(\rho-1)\theta_t]}{1+\sigma E_t}\frac{dE_t}{d\theta_t}\right] > 0, & \text{(A.38)} \\
\frac{dG_t}{d\rho} &= n\tau(w-1)\left[\theta_t\ln(1+\sigma E_t) + \frac{\sigma[1+(\rho-1)\theta_t]}{1+\sigma E_t}\frac{dE_t}{d\rho}\right] > 0, & \text{(A.39)} \\
\frac{dG_t}{dw} &= \tau + \tau(w-1)[1+(\rho-1)\theta_t]\left[\ln(1+\sigma E_t) + \frac{\sigma n}{1+\sigma E_t}\frac{dE_t}{dn}\right] > 0, & \text{(A.40)} \\
\frac{dG_t}{dw} &= n\tau[1+(\rho-1)\theta_t]\left[\ln(1+\sigma E_t) + \frac{\sigma(w-1)}{1+\sigma E_t}\frac{dE_t}{dw}\right] > 0. & \text{(A.41)}
\end{aligned}$$

Proof of Proposition 9

Rewrite (4.10) by using (2.6) as

$$\frac{(1+\alpha)(1+\sigma E)}{\sigma\left\{\frac{(1-\tau)[1+\theta(w-1)]}{n}-\delta-E\right\}}=\frac{\phi[1+(\rho-1)\theta]}{\frac{1}{w-1}+\theta}+\beta[1+(\rho-1)\theta]$$

$$\Rightarrow\quad\frac{(1+\alpha)e^{\frac{\theta}{1+(\rho-1)\theta}}}{\frac{\sigma(1-\tau)[1+\theta(w-1)]}{n}+1-\sigma\delta-e^{\frac{\theta}{1+(\rho-1)\theta}}}=[1+(\rho-1)\theta]\left[\beta+\frac{\phi(w-1)}{1+\theta(w-1)}\right],$$

$$\Rightarrow\quad\frac{\sigma(1-\tau)}{n}=\frac{e^{\frac{\theta}{1+(\rho-1)\theta}}+\sigma\delta-1}{1+\theta(w-1)}+\frac{1+\alpha}{1+(\rho-1)\theta}\cdot\frac{e^{\frac{\theta}{1+(\rho-1)\theta}}}{\beta+(\beta\theta+\phi)(w-1)},\quad\text{(A.42)}$$

which implies that condition (4.11) holds.

Proof of Proposition 10

The welfare-maximizing τ satisfies the following first order condition:

$$\frac{dV(\tau)}{d\tau}=\frac{\phi}{\tau}+\frac{\partial V}{\partial\theta}\frac{d\theta}{d\tau}=0.\quad\text{(A.43)}$$

V is a continuous function of (θ,τ) where $\theta,\tau\in[0,1]$ constitute a compact set. Therefore, the optimal θ and τ must exist.

References

Allingham, Michael, and Agnar Sandmo. 1972. "Income Tax Evasion: A Theoretical Analysis." *Journal of Public Economics*, 1(3-4): 323–338.

Almond, Douglas, Janet Currie, and Valentina Duque. 2018. "Childhood Circumstances and Adult Outcomes: Act II." *Journal of Economic Literature*, 56(4): 1360–1446.

Auerbach, Alan J., and James R. Hines. 2002. "Taxation and Economic Efficiency." *Handbook of Public Economics*, 3: 1347–1421.

Auerbach, Alan J., Jagadeesh Gokhale, and Laurence Kotlikoff. 1994. "Generational Accounting: A Meaningful Way to Evaluate Fiscal Policy." *Journal of Economic Perspectives*, 8(1): 73–94.

Barr, Nicholas. 1992. "Economic Theory and the Welfare State: A Survey and Interpretation." *Journal of Economic Literature*, 30(2): 741–803.

Becker, Gary. 1991. *A Treatise of the Family*. Cambridge, M.A.: Harvard University Press.

Becker, Gary, and H. Gregg Lewis. 1973. "On the Interaction Between the Quantity and Quality of Children." *Journal of Political Economy*, 81(2): S279–S288.

Becker, Gary, Kevin M. Murphy, and Robert Tamura. 1990. "Human Capital, Fertility, and Economic Growth." *Journal of Political Economy*, 98(5): 12–37.

Behrman, Jere R., Andrew D. Foster, Mark R. Rosenzweig, and Prem Vashishtha. 1999. "Women's Schooling, Home Teaching, and Economic Growth." *Journal of Political Economy*, 107(4): 682–714.

Benabou, Roland. 1996. "Heterogeneity, Stratification, and Growth: Macroeconomic Implications of Community Structure and School Finance." *American Economic Review*, 86(3): 584–609.

Boadway, Robin, and Michael Keen. 2000. "Redistribution" in *Handbook of Income Distribution*, A.B. Atkinson and F. Bourguignon (ed.), Vol. 1, Chapter 12, pp. 677–789, New York: Elsevier.

Boadway, Robin, and Maurice Marchand. 1995. "The Use of Public Expenditures for Redistributive Purposes." *Oxford Economic Papers*, 47(1): 45–59.

Brito, Dagobert, Jonathan Hamilton, Steven Slutsky, and Joseph Stiglitz. 1991. "Dynamic Optimal Income Taxation with Government Commitment." *Journal of Public Economics*, 44(1): 15–35.

Broecke, Stijn. 2016. "Do Skills Matter for Wage Inequality?" *IZA World of Labor*, doi: 10.15185/izawol.232

Cremer, Helmuth, Philippe De Donder, and Pierre Pestieau. 2010. "Education and Social Mobility." *International Tax and Public Finance*, 17: 357–377.

Cremer, Helmuth, Firouz Gahvari, and Pierre Pestieau. 2011. "Fertility, Human Capital Accumulation, and the Pension System." *Journal of Public Economics*, 95(11–12): 1272–1279.

Cremer, Helmuth, Firouz Gahvari, and Pierre Pestieau. 2017. "Uncertain Altruism and the Provision of Long Term Care." *Journal of Public Economics*, 151(C): 12–24.

Cremer, Helmuth, Pierre Pestieau, and Jean-Charles Rochet. 2001. "Direct versus Indirect Taxation: the Design of the Tax Structure Revisited." *International Economic Review*, 42: 781–799.

De la Croix, David, and Matthias Doepke. 2003, "Inequality and Growth: Why Differential Fertility Matters." *American Economic Review*, 93(4): 1091–1113.

De la Croix, David, and Matthias Doepke. 2004. "Public versus Private Education when Differential Fertility Matters." *Journal of Development Economics* 73(2): 607–629.

De la Croix, David, and Philippe Michel. 2002. *A Theory of Economic Growth: Dynamics and Policy in Overlapping Generations*. New York, US: Cambridge University Press.

Del Rey, Elena, and Miguel-Angel Lopez-Garcia. 2013. "Optimal Education and Pensions in an Endogenous Growth Model." *Journal of Economic Theory*, 148(4): 1737–1750.

Del Rey, Elena, and Miguel-Angel Lopez-Garcia. 2016. "Endogenous Growth, Education Subsidies, and Intergenerational Transfers." *Economic Modelling*, 52(B): 531–539.

Docquier, Frederic, Oliver Paddison, and Pierre Pestieau. 2007. "Optimal Accumulation in an Endogenous Growth Setting with Human Capital." *Journal of Economic Theory*, 134(1): 361–378.

Ehrlich, Isaac, and Francis T. Lui. 1991. "Intergenerational Trade, Longevity, and Economic Growth." *Journal of Political Economy*, 99(5): 1029–1059.

Fan, C. Simon, and Oded Stark. 2008. "Looking at the 'Population Problem' Through the Prism of Heterogeneity: Welfare and Policy Analyses." *International Economic Review*, 49(3):799–835.

Fan, C. Simon, and Jie Zhang. 2013. "Differential Fertility and Intergenerational Mobility under Private versus Public Education." *Journal of Population Economics*, 26(3): 907–941.

Fleurbaey, Marc. 2008. *Fairness, Responsibility, and Welfare.* Oxford: Oxford University Press.

Fleurbaey, Marc. 2009. "Beyond GDP: The Quest for a Measure of Social Welfare." *Journal of Economic Literature*, 47(4): 1029–1075.

Galor, Oded, and Joseph Zeira. 1993. "Income Distribution and Macroeconomics." *Review of Economic Studies*, 60(1): 35–52.

Galor, Oded, and David N. Weil. 2000. "Population, Technology, and Growth: From Malthusian Stagnation to the Demographic Transition and Beyond." *American Economic Review*, 90(4): 806–828.

Ganelli, Giovanni, and Juha Tervala. 2009. "Can Government Spending Increase Private Consumption? The Role of Complementarity." *Economics Letters*, 103(1): 5–7.

Geller, Andrew M., and Harold Zenick. 2005. "Aging and the Environment: A Research Framework." *Environmental Health Perspectives*, 113(9): 1257–1262.

Glomm, Gerhard, and B. Ravikumar. 1992. "Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality." *Journal of Political Economy*, 100(4): 818–834.

Glomm, Gerhard, and B. Ravikumar. 1997. "Productive government expenditures and long-run growth." *Journal of Economic Dynamics and Control*, 21(1): 183–204.

Hamilton, Jonathan, and Steven Slutsky. 2007. "Optimal Nonlinear Income Taxation with a Finite Population." *Journal of Economic Theory*, 132(1): 548–556.

Hanushek, E. A. 1996. "Measuring Investment in Education." *Journal of Economic Perspectives*, 10(4): 9–30.

Harsanyi, John C. 1955. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility." *Journal of Political Economy*, 63(4): 309–321.

Heckman, James, Rodrigo Pinto, and Peter Savelyev. 2013. "Understanding the Mechanisms Through Which an Influential Early Childhood Program Boosted Adult Outcomes." *American Economic Review*, 103(6): 2052–2086.

Kaganovich, Michael, and Itzhak Zilcha. 1999. "Education, Social Security, and Growth." *Journal of Public Economics*, 71(2): 289–309.

Karlsson, Charlie, and Urban Gråsjö. 2014. "Knowledge Flows, Knowledge Externalities, and Regional Economic Development." *Handbook of Regional Science*, Fischer, M.M., and P. Nijkamp (eds), pp. 413–437. Springer Berlin/Heidelberg.

Lee, Ronald Demos, and Andrew Mason. 2011. *Population Aging and the Generational Economy: A Global Perspective.* Cheltenham, UK: Edward Elgar Publishing.

Lucas, Robert E. Jr. 1988. "On the Mechanics of Economic Development." *Journal of Monetary Economics*, 22(1): 3–42.

Mason, Andrew, and Georges Tapinos. 2000. *Sharing the Wealth: Demographic Change and Economic Transfers Between Generations.* Oxford: Oxford University Press.

Meltzer, Allan H., and Scott F. Richard. 1981. "A Rational Theory of the Size of Government." *Journal of Political Economy*, 89(5): 914–927.

Mirrlees, James A. 1971. "An Exploration in the Theory of Optimum Income Taxation." *Rreview of Economic Studies*, 38(2): 175–208.

Mookherjee, Dilip, and Stefan Napel. 2007. "Intergenerational Mobility and Macroeconomic History Dependence." *Journal of Economic Theory*, 137(1): 49–78.

Musgrave, Richard Abel. 1959. *The Theory of Public Finance*. New York, US: McGraw Hill.

OECD. 2018. General Government Spending (Indicator). doi: 10.1787/a31cbf4d-en (accessed on 29 July 2018)

Pestieau, Pierre, and Mathieu Lefebvre. 2018. *The Welfare State in Europe: Economic and Social Perspectives*, Second Edition, Oxford: Oxford University Press.

Phaneuf, Daniel, and Till Requate. 2017. *A Course in Environmental Economics: Theory, Policy, and Practice*, New York, US: Cambridge University Press.

Piketty, Thomas, and Emmanuel Saez. 2013. "Optimal Labor Income Taxation." *Handbook of Public Economics*, Vol. 5: 391–474.

Possen, Uri, and Steven Slutsky. 1980. "Public-Private Consumption Tradeoffs and the Balanced Budget Multiplier." *Quarterly Journal of Economics*, 95(4): 679–702.

Ramey, Garey, and Valerie A. Ramey. 2010. "The Rug Rat Race." *Brookings Papers on Economic Activity*, 41(1): 129–176.

Rangel, Antonio. 2003. "Forward and Backward Intergenerational Goods: Why Is Social Security Good for the Environment?" *American Economic Review*, 93(3): 813–834.

Rosenthal, Stuart, and William Strange. 2008. "The Attenuation of Human Capital Spillovers." *Journal of Urban Economics*, 64(2): 373–389.

Samuelson, Paul A. 1954. "The Pure Theory of Public Expenditure." *Review of Economics and Statistics*, 36(4): 387–389.

Slemrod, Joel. 2007. "Cheating Ourselves: The Economics of Tax Evasion." *Journal of Economic Perspectives*, 21(1): 25–48.

Stiglitz, Joseph E. 2015. "Leaders and followers: Perspectives on the Nordic model and the economics of innovation." *Journal of Public Economics*, 127: 3–16.

Varvarigos, Dimitrios, and Intan Zanariah Zakaria. 2013. "Endogenous Fertility in a Growth Model with Public and Private Health Expenditures." *Journal of Population Economics*, 26(1): 67–85.

Vernon, Philip E. 1979. *Intelligence: Heredity and Environment*, San Francisco: W. H. Freeman and Company.

Willis, Robert. 1973. "A New Approach to the Economic Theory of Fertility Behavior." *Journal of Political Economy*, 81(2): S14–S64.

World Bank. 2017. "World Development Indicators 2017." Washington, DC: World Bank.