2020/21

DP

Simon Fan, Yu Pang and Pierre Pestieau

Nature versus nurture in social mobility under private and public education systems

CORE
CENTER FOR OPERATIONS RESEARCH AND ECONOMETRICS

CORE
Voie du Roman Pays 34, L1.03.01
B-1348 Louvain-la-Neuve
Tel (32 10) 47 43 04
Email: immaq-library@uclouvain.be
https://uclouvain.be/en/research-institutes/lidam/core/discussion-papers.html

Nature versus Nurture in Social Mobility under Private and Public Education Systems

Simon Fan Yu Pang Pierre Pestieau*

Abstract

This paper analyzes the roles of innate talent versus family background in shaping intergenerational mobility and social welfare under different education systems. We establish an overlapping-generations model in which the allocation of workers between a high-paying skilled labor sector and a low-paying unskilled labor sector depends on talent, parental human capital, and educational resources, and the wage rate of skilled workers is determined by their average talent. Our model suggests that under the private education system, there is a negative relationship between income inequality and social mobility, and the steady-state average talent of skilled workers decreases with educational investments. Under the public education system that provides all children with equal educational resources, the allocation of workforce depends more on talent and less on family background. Consequently, both mobility and inequality increase, and social welfare may improve under reasonable conditions. When private educational investments are allowed on top of public education, the steady-state social welfare increases further. Moreover, if some parents are myopic, public education yields the highest welfare.

JEL Classifications: H20, H31, H50, O11

Keywords: Innate Ability; Private Education; Public Education; Intergenerational Mobility

*Fan: Department of Economics, Lingnan University, Hong Kong (fansimon@ln.edu.hk); Pang: School of Business, Macau University of Science and Technology, Macau (ypang@must.edu.mo); Pestieau (corresponding author): Department of Economics, University of Liège and CORE, Belgium, and Toulouse School of Economics, France (p.pestieau@ulg.ac.be).
The usual disclaimer applies.

1 Introduction

To achieve future career success, is it better to be born rich or to be born intelligent? The answer may depend on the existing education system. The private school system benefits affluent families only. A wealthy parent can offer generous funding for her child's schooling to place him to an elite university and then a high-paying job, even if he is not very smart. The further down the economic ladder a parent is, the less likely is it that her child will attain adequate educational resources. Poor children have a slim chance of escaping poverty not because they lack talent but because material disadvantage holds them back. In contrast, the public school system provides all children with an equal learning opportunity, which helps to mitigate the effects of adverse family backgrounds and to give a heavier weight to innate ability in the determination of educational attainment. Equally gifted children with low-income backgrounds thus stand on the same ground to compete with rich peers, and the gates of upward mobility open for them.

This paper analyzes the roles of individual innate ability ("nature") versus family and school inputs ("nurture") in shaping social mobility and welfare under different educational policies. Intergenerational mobility has received increasing attention in the economic literature. A growing body of empirical studies illustrate its inverse relationship with income inequality, which is known as the "Great Gatsby curve."[1] While income inequality measures the pay gap between rich and poor within a single generation, intergenerational mobility reflects the extent to which children's socioeconomic outcomes differ from those of their parents. Reducing inequality and promoting mobility are both important from the welfare perspective, and the poor tend to resent inequality less if their children are more assured of equal opportunity.[2]

Innate ability and educational expenditure are two major determinants of children's outcome (Becker and Tomes 1986, Abbott, Gallipoli, Meghir, and Violante 2019). People are endowed with different abilities, which may influence their labor productivity and hence income levels (Spence

[1] See, for example, Corak (2013), Chetty, Hendren, Kline, Saez, and Turner (2014), Jerrim and Macmillan (2015), Goldrick-Rab, Kelchen, Harris, and Benson (2016), and Alesina, Stantcheva, and Teso (2018), and Becker, Kominers, Murphy and Spenkuch (2018).

[2] The concern about mobility has a strong philosophical foundation. In his theory of social justice, Rawls (1971) holds that those who have the same talent and the same willingness to use them should have the same prospects of success regardless of their initial situations. Roemer and Trannoy (2016, p. 1289) address that "*equality of opportunity can be described as seeking to offset differences in outcomes attributable to luck, but not those differences in outcomes for which individuals are responsible.*" Also see Piketty (1995), Roemer (1998), Fleurbaey (1995, 2008), Alesina and Angeletos (2005), and Saez and Stantcheva (2016).

1973, Stiglitz 1975).[3] The productivity of human capital intensive industries usually relies on the *average* talent of labor force. For example, a software developer's innovation capability is related more to his innate ability than to his early school performance. Bill Gates and Steve Jobs are two examples of gifted innovators who made pioneering contributions to the development of the IT industry and generated positive externalities for other engineers. The theoretical literature, notably Lucas (1988), postulates the spillover effect of human capital and the importance of average human capital in economic development. In this paper, we follow Lucas's (1988) idea but attend to just one aspect of human capital, namely innate ability.

While the endowments of innate ability are fairly evenly distributed across the whole population (Papageorge and Thom *forthcoming*), this may not be the case for the educational resources received from family and school (as covered in more detail in section 2). Such differences lead educational attainment and career success to be weighted in favor of the haves over the have nots. In the United States, a low-income kindergartner with high test scores for academic talent has only 30% chance of obtaining a college degree and a desirable entry-level job in adulthood, compared to a 70% chance for his rich peer who has low scores (Carnevale, Fasules, Quinn, and Campbell 2019). In addition, 24% of high-potential people born to low-income fathers graduate from college, compared to 63% of those born to high-income fathers, while 27% of low-potential people with high-income fathers graduate from college, which is a greater proportion than that of smartest people from poorest families (Papageorge and Thom *forthcoming*). These marked contrasts imply that family background, and relatedly educational expenditure, play a prominent role in shaping a child's outcome.

We develop an overlapping generations model with a high-paying skilled labor sector and a low-paying unskilled labor sector. A child receives education, and her probability of becoming a skilled worker is governed by her innate ability, parental human capital, and educational resources. Innate ability is exogenously given and is random among the entire population. Human capital can be transmitted between generations in a nonpecuniary manner (Benabou 1996, Becker, Kominers, Murphy, and Spenkuch 2018). We assume that the productivity and wage rate of skilled workers depend on their *average* innate ability to capture the human capital spillover effect between skilled

[3] The literature on the "signaling" and "screening" roles of education in the labor market posits that education can tell the employer about a job candidate's ability, although it contributes little in itself to worker productivity.

workers.

We examine how private and public education systems allocate children with different innate abilities and family backgrounds between the skilled and unskilled labor sectors.[4] Under the private school system, rich children receive more educational investments than poor children not only because their parents have a higher income but also because their parents perform more efficiently in using educational resources. For children with equal ability, those from disadvantaged families are less likely to become high-paid skilled workers in future, which hinders intergenerational mobility. In *laissez-faire* equilibrium, low-ability children with skilled parents have a good chance of getting hired in the skilled labor sector, which undermines the productivity of skilled workers and keeps their wage rate low. Our analysis also indicates that income inequality (as measured by the Gini index) is inversely related to social mobility not only in a given period but also along the path of transitional dynamics, which suggests the existence of the "Great Gatsby curve".

Under the public school system, the government imposes a proportional income tax on workers and uses the revenues to provide every child with equal educational expenditure. The allocation of the labor force thus relies more on innate ability and less on family background, which promotes intergenerational mobility. Compared with private education, public education gives talented people an advantage of being employed in the skilled labor sector, leading to an increase in the average ability of skilled workers and hence their productivity and wage rate. The wider pay gap between skilled and unskilled workers makes income inequalities worse. Overall, the public school system may help to improve welfare under certain conditions. If parents are allowed to make private investments in their children's education in addition to public schooling (a "top-up system"), social welfare will increase further.

We also perform a simulation illustrating the welfare consequences of different educational and fiscal policies. We compare three sets of circumstances, namely the *laissez-faire* case, government intervention in the form of income redistribution ("transfer in cash"), and government intervention in the form of public education ("transfer in kind"). Our simulation finds that transfer in cash leads to lower income inequality and a small improvement in intergenerational mobility, while transfer in kind results in a greater income inequality and a substantial improvement in intergenerational

[4]Earlier literature that compares private with public education includes Glomm and Ravikumar (1992), Epple and Romano (1998), Fernandez and Rogerson (2003), and de la Croix and Doepke (2004, 2009) among others.

mobility. Transfer in cash tends to yield a higher social welfare than transfer in kind. Moreover, we examine how social welfare will change if a fraction of poor parents are myopic and spend little on their children's education.[5] In this case, transfer in kind enhances intergenerational mobility and achieves the highest social welfare.

The remainder of this paper is organized as follows. Section 2 presents some empirical evidence that motivates this paper. Section 3 lays out our model setup. Section 4 examines how three different education systems shape the steady state. Section 5 performs a numerical analysis comparing mobility, inequality, and welfare under various educational and fiscal policies and extends the analysis to consider the existence of myopic parents. Section 6 concludes.

2 Empirical Motivation

This paper is motivated by some striking cases of rich parents spending an tremendous amount of money on their children's education to send them to prestigious universities. The following is a CNN report from 2012.[6]

> *How much would you pay to get your child into an Ivy League university? For Gerard and Lily Chow, it seems the sky was the limit. In 2007, the Hong Kong couple enlisted Harvard-lecturer-turned-admissions-consultant Mark Zimny to steer their two sons through elite U.S. boarding schools into a top-ranked university preferably Harvard. For a monthly $4000 fee per child, their total education management package included extensive admissions counseling, arranging homestays, private tutoring, and extra-curricular activities, whereby Zimny and his team functioned as parents away from parents for their sons. The Chows later switched to a retainer of $1 million per child.*

The urge to secure children's success has even led to an array of university admissions scandals. For example, in May 2019, it was found that the family of a Stanford student had paid $6.5 million to help her get admitted, including bribing university administrators and sports coaches. The student also confessed in a webcast that her IQ is not quite high and her early academic performance

[5]Myopia is a widely recognized topic in public economics literature. See Lindbeck and Persson (2003), Diamond (2004), Cremer, De Donder, Maldonado, and Pestieau (2007), and Cremer and Pestieau (2011) among others.

[6]See "Hong Kong in hot pursuit of Ivy League education by Alexis Lai, CNN.com, 3 December 2012.

was mediocre.[7]

While rich parents are willing to make enormous outlays to secure an academic certificate for their children, poor parents lack the financial resources to compete and find themselves helpless. The wide gap in educational resources between rich and poor manifests itself in at least four ways. First, rich parents tend to buy expensive houses in nice neighborhoods with the top-performing school districts. For example, as rising income inequality has translated into a residential sorting effect for American households with children since 1990, rich and poor have become increasingly unlikely to share the same neighborhoods (Owens 2016). Income inequality has also been identified as a major predictor of income segregation between school districts (Owens, Reardon, and Jencks 2016). Neighborhoods influence children's long-term outcome through childhood exposure effects. Chetty, Hendren, and Katz (2016) show that moving to a low-poverty neighborhood before the teenage years improves the likelihood of college attendance and increases earnings. In a study of more than 7 million families that moved across US counties, Chetty and Hendren (2018) show that the neighborhood in which a child grows up considerably shapes his social mobility.

Second, in societies where free public education plays a limited role, parents have to consider sending their children to private schools, which usually charge very high tuition and accommodation fees. In that case, only wealthy children can go to school, or at least, to a high-quality school. Poor children often have to attend low-quality schools, drop out after a few years of schooling, or even do not attend school at all. Skiba et al. (2008) suggest that poor children are more likely to attend schools with a high rate of teacher turnover, fewer experienced teachers, and larger class sizes. Chetty, Friedman, and Rockoff (2014) demonstrate that a better school with high-quality teachers substantially increases students' educational attainment.

Third, even in developed countries where all children attend primary and middle schools under a free public education system, parents are required to pay for their children's pre-school education. Rich children have access to better early-age educational resources in a high-quality kindergarten where their teachers often receive professional training. It is well established in the literature that early childhood intervention lays the foundation for subsequent academic performance (Shonkoff and Phillips 2000) and plays a powerful role in shaping human capital formation (Heckman, Pinto,

[7] See "Hard work got me into Stanford University, says Chinese student in viral video after parents paid US$6.5 million to get her accepted." By Laurie Chen, *South China Morning Post*, 3 May 2019.

and Savelyev 2013).

Fourth, richer parents tend to spend more on their children's after-school training. The global private tutoring market is booming, with East Asia leading the way. For example, Korea's household expenditure on private tutoring was estimated at 2.8% of GDP in 2006, equivalent to 80% of government spending on primary and secondary school education (Kim and Lee 2010). In Hong Kong, the average tutoring spending for secondary school students in 2010 was 8.7% of household income (Bray et al. 2014). According to the 2009 survey of the Programme for International Student Assessment, about half of students in many western countries used after-school tutoring. In the UK, children from some ultra-rich families take lessons at home with a rotation of tutors who charge £60–100 per hour.[8] Poor families, however, do not have the money to hire private tutors.

3 Model Setup

Consider an overlapping generations economy where a mass of individuals live for two periods – childhood and adulthood. Each child receives education to pursue a college degree but does not make any decision. Every adult bears one child and works for a salary. All adults who successfully obtain a degree become high-paid skilled workers (s), while the others become low-paid unskilled workers (u).

3.1 Preference

There is a population L_t of adults in period t, who are labelled as members of "generation t." A representative member spends her wage, w_t, on private consumption, c_t, and her child's education, e_t, without leaving any bequest. Her budget constraint thus amounts to

$$w_t = c_t + e_t. \tag{1}$$

The representative member's utility, v_t, increases with both consumption, c_t, and the quality of her child, q_t, as given by

$$v_t = \ln c_t + \alpha q_t, \tag{2}$$

[8] See "School's out forever: Why super-rich parents are opting to educate their children using private tutors?" by Joshi Herrmann, *Evening Standard*, 13 February 2015.

where the parameter $\alpha > 0$ represents the preference over the quality of the child.

Denote θ_t as the proportion of skilled workers in generation t. In the beginning of period t, a member expects to obtain v_t^s with a probability of θ_t and v_t^u with a probability of $1-\theta_t$. In line with Harsanyi (1955), social welfare in period t, V_t, is defined as the individual expected utility:

$$V_t = \theta_t v_t^s + (1-\theta_t) v_t^u. \tag{3}$$

3.2 Skill and Ability

The quality of a child is measured by the probability that she receives a degree and becomes a skilled worker in her adulthood. A simply formulation of $q_t^i \in [0,1]$, where $i \in \{s, u\}$ denotes the parent type, is

$$q_t^i = q(a_{t+1}, e_t^i) = \ln(1 + a_{t+1}) + \beta^i \ln e_t^i, \tag{4}$$

which increases with the child's innate ability, a_{t+1}, and the received educational spending, $e_t^i \geqslant 1$, and exhibits decreasing returns to both arguments. The parameter β^i measures the relative weight of educational spending in determining the quality of the child and also reflects to what extent the prevailing labor market and education system are meritocratic. Assume that parents with a college degree perform no worse at using educational resources than those who have no degree, namely $\beta^s \geqslant \beta^u > 0$.[9]

Suppose that an individual's innate ability is unrelated to the family she was born into, and it is not revealed in childhood and remains publicly unobservable in adulthood. For simplicity, a_t is a binary random variable satisfying

$$a_t = \begin{cases} a^H > 0 & \text{with a probability of } \lambda \\ a^L = 0 & \text{with a probability of } 1-\lambda \end{cases}, \tag{5}$$

where $\lambda \in (0,1)$ is exogenous. Equation (5) shows that an individual may be endowed with high ability, a^H, or low ability, a^L, which is normalized to zero. By law of large numbers, generation t is composed of a fraction λ of high ability members and a fraction $1-\lambda$ of low ability members.

[9] For example, educated parents tend to be better at encouraging and motivating their children, have a more positive attitude toward learning, own more books and read regularly, and provide healthier diets (e.g., Papageorge and Thom, *forthcoming*).

Rewrite (4) by using (5):

$$
\begin{aligned}
q_t^i &= \lambda\Big[\ln(1+a^H) + \beta^i \ln e_t^i\Big] + (1-\lambda)\Big[\ln(1+a^L) + \beta^i \ln e_t^i\Big] \\
&= \lambda \ln\big(1+a^H\big) + \beta^i \ln e_t^i, \qquad (6)
\end{aligned}
$$

where the first term captures the role of "nature" in shaping child outcome, while the second term reflects the effect of "nurture" (including family environment and school input).

Following Iyigun (1999), we measure the intergenerational social mobility by the odds ratio – the probability that children of unskilled workers become skilled relative to the probability that children of skilled workers become skilled. Accordingly, we propose the definition as below:

Definition 1 *Intergenerational social mobility in period t is measured by*

$$
M_t := \frac{q_t^u}{q_t^s} = \frac{\lambda \ln(1+a^H) + \beta^u \ln e_t^u}{\lambda \ln(1+a^H) + \beta^s \ln e_t^s}. \qquad (7)
$$

Because $\beta^u \leqslant \beta^s$ by assumption and also $e^u \leqslant e^s$ in equilibrium, we have $M_t \in (0,1)$. When M gets close to one, children from different families have an equal opportunity of becoming skilled in future. If parents differ more significantly in their educational spending and the efficiency of using educational resources, then M will be lowered, indicating that the prospect for convergence tends to be constrained. As M approaches zero, almost all skilled workers have a skilled parent.

3.3 Production and Wage

There are two sectors in the economy – the unskilled and skilled labor sectors – which produce the unique final good competitively using their own type of labor with a constant-returns-to-scale technology. Let Y_t^u and Y_t^s be the output in the unskilled and skilled labor sectors in period t and L_t^u and L_t^s be their labor inputs. Their production functions are specified as follows:

$$
Y_t^u = \phi L_t^u \qquad\qquad Y_t^s = f(\bar{a}_t) L_t^s, \qquad (8)
$$

where $\phi > 0$ represents the productivity of unskilled workers, which is independent of their ability and remains constant over time. The productivity of skilled workers, $f(\bar{a}_t)$, is a function of their average innate ability $\bar{a}_t$, which in turn can be written as

$$\bar{a}_t = \mu_t a^H, \tag{9}$$

where μ_t is the share of the high ability in skilled workers of generation t. Equation (9) essentially means that only high ability workers contribute to the sectoral productivity.[10]

With the price of final goods in every period being normalized to unity, it is straightforward to determine that the equilibrium wage of an unskilled worker is $w_t^u = \phi$ under perfect competition. An unskilled worker earns a wage $w_t^s = f(\bar{a}_t)$ in equilibrium.[11] Assume that a skilled worker is more productive and is paid a higher wage rate than an unskilled worker:[12]

$$f(\bar{a}_t) = \delta \bar{a}_t > \phi > 0, \tag{10}$$

where the parameter $\delta > 0$ measures the marginal return to skilled workers' average innate ability.

3.4 Population Composition

We proceed to discuss demographic characteristics from the macroeconomic perspective. The population of generation t, L_t, can be decomposed with respect to labor skill into two cohorts:

$$L_t^s = \theta_t L_t, \qquad\qquad L_t^u = (1-\theta_t)L_t. \tag{11}$$

Skilled and unskilled workers expect to have a number of $q_t^s L_t^s$ and $q_t^u L_t^u$, respectively, of children who will become skilled workers in future. In generation t, the proportion of skilled workers is $\theta_{t+1} = (q_t^s L_t^s + q_t^u L_t^u)/L_{t+1}$. Given that the fertility rate is one (i.e., $L_{t+1} = L_t$ for all t), we can rewrite this proportion by using (6) and (11):

[10] This result reflects the empirical observations that talented people, as opposed to the "trivial many", are making a big difference and are the main driver of a company's success. See "Talent matters even more than people think" by Tomas Chamorro-Premuzic, *Harvard Busiess Review*, 4 October 2016.

[11] In our stylized model, individual talent is not rewarded in the labor market. An alternative assumption would be that a skilled worker's wage depends on both her own talent and the average talent, but this would complicate analysis without a qualitative change in our results under transitional dynamics. The steady states under such an assumption would be hard to derive, since it would imply that the individual heterogeneity would increase over time.

[12] Skilled workers earn a higher wage than unskilled workers regardless of their intelligence. As Borghans, Golsteyn, Heckman, and Humphries (2016) find, financial success and high innate talent have little correlation.

$$\theta_{t+1} = \lambda \ln(1 + a^H) + \beta^s \theta_t \ln e_t^s + \beta^u (1 - \theta_t) \ln e_t^u, \tag{12}$$

which governs the evolutionary path of θ_t.

By (4), (5), and (11), skilled workers of generation $t + 1$ can be further categorized into four groups whose population sizes are as follows

High ability from a rich family:	$\lambda q(a^H, e_t^s) L_t^s$;
Low ability from a rich family:	$(1 - \lambda) q(a^L, e_t^s) L_t^s$;
High ability from a poor family:	$\lambda q(a^H, e_t^u) L_t^u$;
Low ability from a poor family:	$(1 - \lambda) q(a^L, e_t^u) L_t^u$.

The proportion of the high ability in skilled workers, μ_{t+1}, can be expressed by

$$\begin{aligned} \mu_{t+1} &= \frac{\lambda q(a^H, e_t^s) L_t^s + \lambda q(a^H, e_t^u) L_t^u}{\theta_{t+1} L_{t+1}} \\ &= \frac{\lambda \left[\ln(1 + a^H) + \beta^s \ln e_t^s\right] \theta_t + \lambda \left[\ln(1 + a^H) + \beta^u \ln e_t^u\right] (1 - \theta_t)}{\theta_{t+1}} \\ &= \lambda \left[1 + \frac{(1 - \lambda) \ln(1 + a^H)}{\theta_{t+1}}\right] > \lambda, \end{aligned} \tag{13}$$

which implies that the majority of high ability individuals work in the skilled labor sector. Combining (9) and (13) derives the average innate ability of skilled workers of generation $t + 1$, which is higher than the average innate ability of the whole population:

$$\bar{a}_{t+1} = \lambda a^H \left[1 + \frac{(1 - \lambda) \ln(1 + a^H)}{\theta_{t+1}}\right] > \lambda a^H. \tag{14}$$

4 Equilibrium under Different Educational Systems

In this section, we characterize the educational spending per child (e_t^i), the proportion of skilled workers (θ_t), the average ability of skilled workers ($\bar{a}_t$), and intergenerational mobility (M_t) under three education systems. We first characterize the *laissez-faire* equilibrium outcome – in the absence of public schools, all parents send their children to private schools and optimally choose their own educational investments. Next, we consider that the government funds public schools by imposing a uniform income tax rate on all workers, while parents do not make extra educational investments in their children. Finally, we accommodate private and public education in a joint framework, with parents allowed to decide on private after-school education for their children given that all children are formally educated in public schools.

4.1 Private Education (*Laissez-faire*)

We start with the case where people make decentralized decisions on their educational spending without any government intervention. Taking into account the budget constraint (1) and the quality of child (6), every i-type member of generation t obtains a utility at the level of

$$v_t^i = \alpha\lambda \ln(1 + a^H) + \ln(w_t^i - e_t^i) + \alpha\beta^i \ln e_t^i, \tag{15}$$

where $i \in \{s, u\}$. She chooses educational spending to maximize her utility according to $\frac{\partial v_t^i}{\partial e_t^i} = 0$, which can be solved as

$$e_t^i = \gamma^i w_t^i \qquad \text{where} \quad \gamma^i \equiv \frac{\alpha\beta^i}{\alpha\beta^i + 1} \in (0, 1).$$

Clearly, an i-type worker allocates a fraction γ^i of her wage to her child's education in equilibrium, where $\gamma^s \geqslant \gamma^u$ under the assumption of $\beta^s \geqslant \beta^u$. The optimal choices of a skilled worker and an unskilled worker hence satisfy

$$e_t^s = \gamma^s \delta \bar{a}_t \qquad\qquad e_t^u = \gamma^u \phi. \tag{16}$$

It follows that $e_t^s > e_t^u$: a skilled worker pays a premium for her child's education not only because she receives a higher income ($\delta\bar{a} > \phi$) but also because she is more efficient in using educational resources ($\gamma^s \geqslant \gamma^u$).

The next lemma follows directly from (6) and (16):

Lemma 1 *Under the private school system, a skilled parent is more likely to have a skilled child than an unskilled parent.*

Lemma 1 suggests that a child is more likely to become skilled if her parent is a skilled worker. As equation (6) demonstrates, a rich child's advantage comes from two sources. First, she receives more educational resources from her parent ($e^s > e^u$). Second, her parent performs at least as well as a poor parent in using these resources ($\beta^s > \beta^u$). The intergenerational transmission of human capital and labor earnings is advocated in earlier theories (e.g., Becker and Tomes 1979, 1986) and is supported by empirical examinations (e.g., Behrman and Rosenzweig 2002, Schneider, Hastings, and LaBriola 2018).

We now relate intergenerational mobility with income inequality. Inserting (16) into (7) rewrites intergenerational mobility in period t as

$$M_t = \frac{\lambda \ln(1 + a^H) + \beta^u \ln(\gamma^u \phi)}{\lambda \ln(1 + a^H) + \beta^s \ln(\gamma^s \delta \bar{a}_t)}. \tag{17}$$

At the aggregate level, the skilled and unskilled of generation t receive a payroll of $\delta \bar{a}_t L_t^s$ and ϕL_t^u, respectively. To measure income inequality in period t, we rely on the Gini coefficient, G_t, which can be computed as the share of high-paid workers' income in total income minus their proportion in the working-age population:

$$G_t = \frac{\delta \bar{a}_t L_t^s}{\delta \bar{a}_t L_t^s + \phi L_t^u} - \theta_t = \theta_t \left[\frac{1}{\theta_t + (1 - \theta_t)\phi/(\delta \bar{a}_t)} - 1 \right], \tag{18}$$

which is a function of (ϕ, δ, $\bar{a}_t$, θ_t).

We can see from (17) and (18) that common factors influencing both M_t and G_t are (ϕ, δ, $\bar{a}_t$). An increase in unskilled workers' wage rate (higher ϕ) promotes social mobility and remedies income inequality (higher M_t and lower G_t), and a fall in skilled workers' wage rate (lower δ or $\bar{a}_t$) achieves the same result. In sum, M_t and G_t are inversely correlated. We develop the following proposition to summarize the result:

Proposition 1 *In period t, a lower income inequality (smaller G_t) is associated with a greater intergenerational social mobility (larger M_t) under the private school system.*

Proposition 1 formalizes the important conceptual link between economic inequality and social mobility in a certain period, namely that more equal societies tend to be more mobile. All else being equal, narrowing the wage gap (i.e., small $\delta \bar{a}_t$ and/or large ϕ) helps to mitigate *intra-*

generational earnings inequality. The income effect suggests that the rich will then tend to spend less on education, with the result that their children are less likely to become rich in the future, but the poor will do the opposite, resulting in more opportunities for their children.

Our prediction is consistent with the empirically observed Great Gatsby curve. Recent research based on cross-sectional data finds that higher inequality in childhood is related with lower mobility in adulthood in Sweden (Brandén 2019) and Latin America (Neidhöfer 2019). It is also consistent with the empirical finding that educational attainment mediates the link between social origin and destination (Jerrim and Macmillan 2015).

We proceed to characterize the evolution of (θ_t, $\bar{a}_t$). Plugging (16) into (12) and (14) obtains

$$\theta_{t+1}(\bar{a}_t, \theta_t) = \lambda \ln(1 + a^H) + \beta^s \theta_t \ln(\gamma^s \delta \bar{a}_t) + \beta^u (1 - \theta_t) \ln(\gamma^u \phi) \tag{19}$$

$$\bar{a}_{t+1}(\bar{a}_t, \theta_t) = \left[1 + \frac{(1-\lambda)\ln(1 + a^H)}{\lambda \ln(1 + a^H) + \beta^s \theta_t \ln(\gamma^s \delta \bar{a}_t) + \beta^u (1 - \theta_t) \ln(\gamma^u \phi)}\right] \lambda a^H. \tag{20}$$

These two equations show that θ_t and $\bar{a}_t$ are both path-dependent, namely θ_{t+1} is determined by θ_t, and $\bar{a}_{t+1}$ is determined by $\bar{a}_t$. Define the steady state of the economy as that both θ_t and $\bar{a}_t$ achieve their time-invariant levels, θ and $\bar{a}$.

Definition 2 *The economy reaches the steady state when $\theta_t = \theta$ and $\bar{a}_t = \bar{a}$.*

We characterize the steady-state solutions in the following proposition:

Proposition 2 *Under the private school system, the steady state of the economy is determined by the following two equations*

$$\left[1 + \frac{\beta^u \ln(\gamma^u \phi)}{\lambda \ln(1 + a^H)}\right] \frac{\bar{a}}{a^H} + \beta^s (1 - \lambda) \ln(\gamma^s \delta \bar{a}) = 1 + \left[1 - \lambda + \frac{1}{\ln(1 + a^H)}\right] \beta^u \ln(\gamma^u \phi), \tag{21}$$

$$\theta = \frac{\lambda(1-\lambda) a^H \ln(1 + a^H)}{\bar{a} - \lambda a^H}. \tag{22}$$

Proof. See Appendix. ■

Equation (21) demonstrates that the average innate ability of skilled workers in the steady state, $\bar{a}$, is governed by six parameters, namely $\bar{a}(a^H, \lambda, \alpha, \beta^s, \beta^u, \delta, \phi)$. Inserting $\bar{a}$ solved in equation (21) into equation (22) derives the solution to θ. As equation (22) shows, holding (a^H, λ) constant, θ and $\bar{a}$ are inversely correlated in the steady state. Since a skilled worker is more likely to be high ability than an unskilled worker on average (equation (13)), the average talent of skilled workers tends to decline when workers are increasingly moved from the unskilled to the skilled labor sector.

We proceed to examine the comparative statics of $\bar{a}$ in the following proposition:

Proposition 3 *Under the private school system, the steady-state average innate ability of skilled workers $\bar{a}$: (i) increases with a^H and λ, (ii) decreases with α, (iii) decreases with β^s and β^u, and (iv) decreases with δ and ϕ.*

Proof. See Appendix. ■

Proposition 3(i) is straightforward: all else being equal, if talented people have a higher level of talent and account for a larger population size, then the average talent of skilled workers will be higher in the steady state. The other parts of Proposition 3 can be interpreted from the perspective of parents' educational investments. A parent tends to invest more in her child's schooling if she is more concerned about the educational attainment of her child (larger α), more efficient in using educational resources (larger β^s or β^u), or earns a higher wage (larger δ or ϕ). By (4), educational expenditure and innate ability are substitutes in governing an individual's quality and skill. As a rise in educational spending mitigates the relative weight of natural ability in occupational choice, more mediocre people born into rich families are employed in the skilled labor sector, which undermines the productivity of skilled occupation.

In our model, since natural talent is usually exogenously given, children from poor and rich families tend to have the same probability of having high innate ability. Therefore, an economy becomes more efficient if intergenerational mobility is higher, which will lead to a higher level of innate ability among skilled workers. The next question that arises is how to enhance the efficiency and equity of society by reallocating educational resources across families. This goal can be achieved by taxing the rich and subsidizing the poor. The income effect means that after income redistribution, the rich will spend less on their children's education whereas the poor will spend more. Another solution would be the provision of public education, which aims to provide all children with an equal amount of educational spending. In the next subsections, we investigate the equilibrium outcome when the government finances public education by levying an income tax.

4.2 (Pure) Public Education

In this section, we examine the case in which the government funds public education by collecting income taxes. Suppose that the government imposes a proportional income tax on all workers and spends the revenue on public education such that each child receives equal educational spending. Public education then gives rise to two benefits. First, rich families give pecuniary support to poor families through in-kind redistribution. Second, public education helps to eliminate the gap in received educational resources between rich and poor children, which helps to promote social mobility. We aim to provide a normative analysis of the optimal public education provision.

Consider that the government levies an income tax at the rate of $\tau_t \in (0,1)$ and provides each child with public educational spending e_t^P in period t. The balanced government budget implies:

$$e_t^P = \frac{\tau_t(w_t^s L_t^s) + \tau_t(w_t^u L_t^u)}{L_{t+1}} = \tau_t[\theta_t \delta \bar{a}_t + (1-\theta_t)\phi], \tag{23}$$

which strictly increases with τ_t. By (23), we rewrite the social welfare in (3) as

$$\begin{aligned} V_t &= \theta_t \Big\{ \ln[(1-\tau_t)w_t^s] + \alpha \left[\lambda \ln(1+a^H) + \beta^s \ln e_t^P\right] \Big\} \\ &\quad +(1-\theta_t)\Big\{ \ln[(1-\tau_t)w_t^u] + \alpha \left[\lambda \ln(1+a^H) + \beta^u \ln e_t^P\right] \Big\} \\ &= \ln(1-\tau_t) + \alpha[\beta^s\theta_t + \beta^u(1-\theta_t)]\big\{ \ln\tau_t + \ln[\theta_t\delta\bar{a}_t + (1-\theta_t)\phi]\big\} \\ &\quad +\theta_t \ln(\delta\bar{a}_t) + (1-\theta_t)\ln\phi + \alpha\lambda\ln(1+a^H). \end{aligned} \tag{24}$$

Given $(\theta_t, \bar{a}_t)$, differentiating V_t in (24) with respect to τ_t and setting it to zero yields

$$\frac{dV_t}{d\tau_t} = \frac{\alpha[\beta^s\theta_t + \beta^u(1-\theta_t)]}{\tau_t} - \frac{1}{1-\tau_t} = 0.$$

Denote $\hat{\tau}_t$ the solution to the above equation, where

$$\hat{\tau}_t = 1 - \frac{1}{\alpha[\beta^s\theta_t + \beta^u(1-\theta_t)] + 1}. \tag{25}$$

We have two comments on equation (25). First, the interior solution to the optimal tax rate always falls in the range $(0,1)$. Second, the government will levy a higher tax in period t if people concern more about the quality of child (larger α) and use educational resources more efficiently (larger β^s and β^u) and if skilled workers account for a greater share of working-age population (larger θ_t).

Substituting (25) into (23) and then using (14) yields e_t^P as a function of θ_t:

$$e_t^P = \frac{\lambda\delta a^H[(1-\lambda)\ln(1+a^H)+\theta_t]+(1-\theta_t)\phi}{\{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)]\}^{-1}+1}. \tag{26}$$

We develop the following lemma to compare the private and public school systems:

Lemma 2 *Compared with the private school system, the public school system leads rich children to receive less education and poor children more, i.e.,* $e_t^s > e_t^P > e_t^u$. *Consequently, intergenerational social mobility improves.*

Proof. See Appendix. ■

If the private education system is replaced by a public one, then educational resources tend to gravitate from rich to poor children until each child receives an equal amount. In equilibrium, more talented poor children attain a college degree, crowding out some mediocre wealthy children. It follows that the likelihood of the poor of becoming rich moves closer to that for the rich. The public school system creates a more equal society where a person's career success depends more on innate ability and less on family background.[13] Lemma 2 not only provides a rationale for the wide adoption of public education but also finds some empirical support. For example, Neidhöfer (2019) shows that public education is significantly and positively associated with intergenerational mobility in Latin American countries.

Finally, we examine the steady state of the economy in the following proposition:

Proposition 4 *Under the public school system, the steady state of the economy is determined by*

$$\theta - [\beta^s\theta+\beta^u(1-\theta)]\ln\left\{\frac{\lambda\delta a^H[(1-\lambda)\ln(1+a^H)+\theta]+(1-\theta)\phi}{\{\alpha[\beta^s\theta+\beta^u(1-\theta)]\}^{-1}+1}\right\} = \lambda\ln(1+a^H), \tag{27}$$

$$\bar{a} = \lambda a^H\left[1+\frac{(1-\lambda)\ln(1+a^H)}{\theta}\right]. \tag{28}$$

[13] Hassler and Rodríguez Mora (2000) present a similar result in a different model setup. They suggest that when the world changes slowly, children of skilled workers have an informational advantage for skilled occupations over other children, but if the world changes a great deal between generations, then parents' information becomes less valuable and innate ability becomes more important in social selection. Rather than discuss the environment of economic growth, our paper addresses the role of the education system.

Proof. See Appendix. ■

Equation (27) demonstrates that the steady-state proportion of skilled workers, θ, is governed by $(a^H, \lambda, \alpha, \beta^s, \beta^u, \delta, \phi)$. Consequently, it can be inferred from (28) that the average innate ability of skilled workers in the steady state is also influenced by the seven parameters. Furthermore, by (26), e^P is a constant in the steady state when $(\theta, \bar{a})$ remain constant.

Let us close this subsection by analyzing the comparative static of $\bar{a}$. To simplify the algebra, we focus on a special case of perfect meritocracy (i.e., $\beta^s = \beta^u$).

Proposition 5 *Suppose $\beta^i = \beta$. Under the public school system, the steady-state average ability of skilled workers, $\bar{a}$, decreases with (α, δ, ϕ) if and only if*

$$\beta < \frac{(\delta\bar{a} - \phi)\theta + \phi}{\lambda a^H - \phi}. \tag{29}$$

Proof. See Appendix. ■

Proposition 5 presents a result similar to that in Proposition 3, with the only difference being that condition (29) is imposed in Proposition 5. This condition emphasizes that parents' efficiency in using educational expenditure should be sufficiently low (small β). Parents who care little about the quality of their children (small α) or earn a low income (small δ and ϕ) tend to spend a small amount of money on their children's education. Limited educational resources, coupled with low efficiency in using educational resources, will make innate ability a predominant determinant of one's placement. Intelligent children are thus very likely to become skilled workers, which in turn increases the average ability of the skilled labor sector.

4.3 Topping Up the Public School System

In the previous subsection, we consider that the government optimizes children's education and parents do not make any decision. We now modify the setting by allowing parents to provide their own children, who receive formal education at public schools, with supplementary remedial classes. In period t, the interaction between the government and workers proceeds in a two-stage game. The government moves first to fund public education by imposing the welfare-maximizing income tax, and then workers choose to provide their children with the optimal educational investments.

After-school private tutoring, which is often referred to as "shadow education," is an example of programs existing alongside the formal education system with a growing relative size. But tutoring is often found to be less effective than mainstream schooling. For example, after-school tutoring often takes place off campus; to participate in private tutoring therefore incurs a fixed adjustment cost, such as the time and expense of commuting between the tutoring center and school (Fashola 2001). Besides, the curriculum emphasized by tutoring institutions may not match well with that in schools (Bray 2003). In our model, we consider the discount in private tutoring effectiveness (relative to public education).

In period t, the government finances public education in accordance with its budget constraint (23). A skilled worker has an incentive to send her child to after-school programs, where a tuition fee of e_t^A is charged, so that the child receives an education investment at the aggregate level of

$$e_t^s = e_t^P + \kappa e_t^A, \tag{30}$$

where e_t^P is expressed in (23) and the parameter $\kappa \in (0, 1)$ measures the efficiency loss of tutoring.

A skilled worker's utility can be written as

$$v_t^s = \alpha\lambda \ln(1 + a^H) + \ln\left[(1 - \tau_t)\delta\bar{a}_t - e_t^A\right] + \alpha\beta^s \ln\left(e_t^P + \kappa e_t^A\right).$$

Given the tax rate τ_t and the public educational expenditure per child e_t^P, a skilled worker optimally chooses private educational expenditure to maximize her utility. Taking the first order condition of the above equation with respect to e_t^A and rearranging obtains

$$e_t^A = \gamma^s(1 - \tau_t)\delta\bar{a}_t - \frac{e_t^P}{(1 + \alpha\beta^s)\kappa}. \tag{31}$$

We see a net substitution effect between public education and private tutoring (i.e., e_t^A and e_t^P are inversely related). By (30) and (31), the educational investment received by a rich child is

$$e_t^s = e_t^P + \kappa\left[\gamma^s(1 - \tau_t)\delta\bar{a}_t - \frac{e_t^P}{(1 + \alpha\beta^s)\kappa}\right] = \gamma^s\left[\kappa(1 - \tau_t)\delta\bar{a}_t - e_t^P\right]. \tag{32}$$

In light of discount of private tutoring (i.e., $\kappa < 1$), the government will choose sufficiently large public educational expenditure in equilibrium so that a poor parent needs not to spend on tutoring.[14] In other words, a poor child in period t receives public education only (i.e., $e_t^u = e_t^P$).

In period t, the government chooses the tax rate to maximize social welfare as expressed by

$$\begin{aligned}
V_t &= \theta_t \left\{ \ln\left[(1-\tau_t)w_t^s - e_t^A\right] + \alpha\lambda\ln(1+a^H) + \alpha\beta^s \ln\left[\gamma^s\kappa(1-\tau_t)\delta\bar{a}_t - \gamma^s e_t^P\right]\right\} \\
&\quad + (1-\theta_t)\left\{\ln[(1-\tau_t)w_t^u] + \alpha\left[\lambda\ln(1+a^H) + \beta^u \ln e_t^P\right]\right\} \\
&= \alpha\lambda\ln(1+a^H) + \theta_t \ln\left\{(1-\tau_t)(1-\gamma^s)\delta\bar{a}_t + \frac{\tau_t[\theta_t\delta\bar{a}_t + (1-\theta_t)\phi]}{(1+\alpha\beta^s)\kappa}\right\} \\
&\quad + \theta_t\alpha\beta^s\ln\gamma^s + \theta_t\alpha\beta^s\ln\left\{\kappa(1-\tau_t)\delta\bar{a}_t - \tau_t\left[\theta_t\delta\bar{a}_t + (1-\theta_t)\phi\right]\right\} + (1-\theta_t)\ln\phi \\
&\quad + (1-\theta_t)\ln(1-\tau_t) + (1-\theta_t)\alpha\beta^u\ln[\theta_t\delta\bar{a}_t + (1-\theta_t)\phi] + (1-\theta_t)\alpha\beta^u\ln\tau_t. \qquad (33)
\end{aligned}$$

Since V_t is continuous function of τ_t, which belongs to a compact set $[0,1]$, the optimal solution to τ_t must exist. We denote τ_t^* as the optimal tax rate and propose the next lemma:

Lemma 3 *The top-up education system is operative in period t if and only if*

$$\alpha\beta^s\kappa\left(\frac{1}{\tau_t^*} - 1\right) > \theta_t + (1-\theta_t)\frac{\phi}{\delta\bar{a}_t}. \qquad (34)$$

Proof. See Appendix. ■

Lemma 3 characterizes a sufficient and necessary condition for tutoring classes to operate (i.e., $e^A > 0$). Condition (34) is more likely to hold if the rich concern greatly about the quality of their children (large α) and are efficient in using educational resources (large β^s) and if private tutoring is efficient (large κ). Moreover, this condition is more likely to hold in period t if the government provides a small educational fund (small e_t^P and thus small τ_t^*), the income gap is wide (small $\frac{\phi}{\delta\bar{a}_t}$), and skilled workers account for a small fraction in the working-age population (small θ_t).

Based on Lemma 3, we compare the top-up education system with the other two school systems in the next proposition, which obtains two results:

Proposition 6

(i) Social mobility is higher under the top-up system than under the private school system.

[14] This outcome has a clear empirical counterpart: it is found that family socioeconomic status is positively related with student participation in private tutoring in South Korea (Kim and Lee 2010) and Hong Kong (Bray et al. 2014).

(ii) Holding public expenditure constant (for any given e_t^P), social mobility is the highest under pure public school system.

Proof. See Appendix. ■

Proposition 6 extends Lemma 2 to compare intergenerational social mobility under three education systems. Compared with the private school system, the top-up system lowers the educational investment received by a rich child and increases that received by a poor child. This result follows directly from the income effect: the top-up system involves an income redistributive process that effectively transfers wealth from rich to poor families. Consequently, a rich child has a lower probability of becoming rich in the future while a poor child faces a higher probability, which mitigates the transmission of earnings inequality across generations.

We next compare the pure public and the top-up education systems. Holding public expenditure constant (for any given e_t^P), we can show that social mobility is higher under the pure public school system. Nevertheless, the equilibrium public educational spending per child is also larger under the pure public school system for two reasons. First, as skilled workers make additional educational investments (i.e., $e^A > 0$) under the top-up system, they will afford a lower tax to finance public education ("income effect"). Second, as skilled workers have already provided their children with private education under the top-up system, the marginal benefit of further educating their children in public schools is lower ("substitution effect").

5 A Numerical Comparison of Different Education Systems

To compare the equilibrium outcome of the three education systems, we rely on a simple quantitative example in this section. We conduct three sets of simulations – (i) a dynamic Great Gatsby curve under the private school system, (ii) the steady-state socioeconomic features under different systems, and (iii) social welfare in transitional dynamics and steady state under different systems. Our simulations are performed based on the parameter values given in Table 1.

Figure 1 charts income inequality versus social mobility over time under *laissez-faire*. In the initial period (period 0), the Gini coefficient is $G_0 \approx 0.302$ while mobility is $M_0 \approx 0.458$. In the next period, the income distribution becomes more even as the Gini coefficient falls to $G_1 \approx 0.247$; meanwhile, intergenerational mobility increases to $M_1 \approx 0.495$. An intertemporal comparison

Table 1. Benchmark Values for Key Parameters

Parameter	Description	Value
a^H	Level of high ability	3
λ	Probability that an individual is endowed with high ability	0.2
α	Preference over the quality of the child	5
β^s	Educational efficiency of rich families	0.2
β^u	Educational efficiency of poor families	0.1
δ	Impact of average ability on skilled labor productivity	5
ϕ	Unskilled labor productivity	4
θ_0	The initial fraction of skilled workers	0.3
κ	Discount of after-school private education	0.9

indicates "more equality, more mobility." The arrows depict the direction of motion: the economy sees continuous and simultaneous improvements in equality and mobility until it reaches the steady state ($G \approx 0.235$, $M \approx 0.502$). In short, Figure 1 shows a Great Gatsby curve along the *dynamic* path, which complements Proposition 1 that implies a *static* inverse relationship between economic inequality and social mobility under *laissez-faire*.

Figure 1: The Great Gatsby Curve along the Dynamic Path under the Private School System

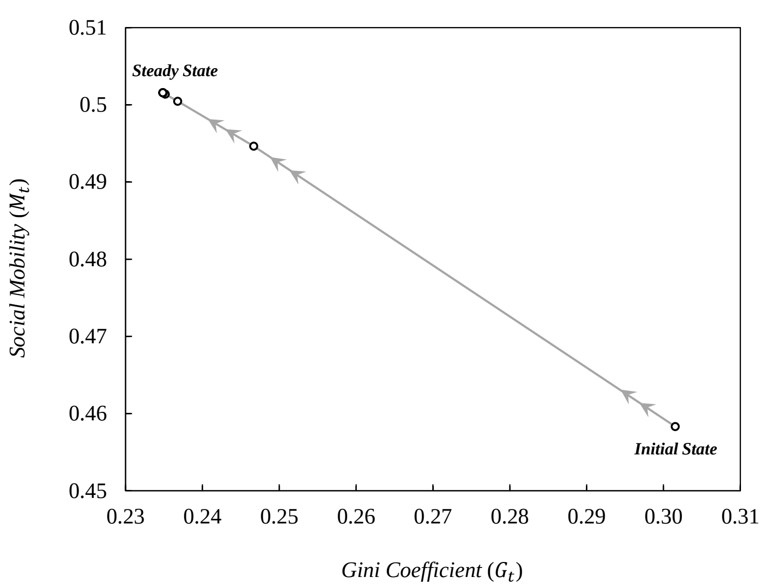


Table 2 reports the simulated outcomes of (θ, $\bar{a}$, M, G) in cases where the formal education is private (in the first panel) and where it is public (in the second panel). For the private school system, we consider three scenarios: [1] *laissez-faire* (section 4.1), [2] the government imposes an income tax at the same rate of $\hat{\tau}_t$ as in section 4.2 and returns the revenue equally to workers, and [3] the government imposes an income tax at the rate of τ_t^* as in section 4.3 and returns the revenue equally to workers. Comparing [1] with [2] and [3] captures the effect of government interference via "transfer in cash". For the public school system, we probe two scenarios: one in which extra private educational investment is not allowed (section 4.2) and one in which it is allowed (section 4.3). Comparing [2] and [3] with [4] and [5] shows the different impacts of "transfer in cash" and "transfer in kind".

We see from the first column of Table 2 that under the private school system, more workers are hired in the skilled labor sector (larger θ) after transfer in cash. The immediate policy implication is that income redistribution is a useful instrument to help more children to obtain a college degree. Besides, the steady-state proportion of skilled workers is smallest (0.4280) when the government provides pure public education but is largest (0.4431) when top-up is allowed. A possible explanation is that under the parameter configuration in Table 1, children receive more educational resources in general when public education is coupled with private education. Note that there is a tradeoff between $\bar{a}$ and θ in the steady state by equation (22). Holding other things fixed, income redistribution tends to decrease the average ability of skilled workers. Public schooling seems the most effective way of placing smart people into the skilled labor sector.

Table 2. The Steady-state Properties Under Different Systems

	θ	$\bar{a}$	M	G
Private School System				
[1] Laissez-faire	0.4398	2.1131	0.5016	0.2349
[2] Transfer-in-cash with Tax Rate $\hat{\tau}_t$	0.4410	2.1089	0.5745	0.1366
[3] Transfer-in-cash with Tax Rate τ_t^*	0.4413	2.1080	0.5492	0.1705
Public School System				
[4] No Top-up	0.4280	2.1547	0.7838	0.2404
[5] With Top-up	0.4431	2.1019	0.5937	0.2333

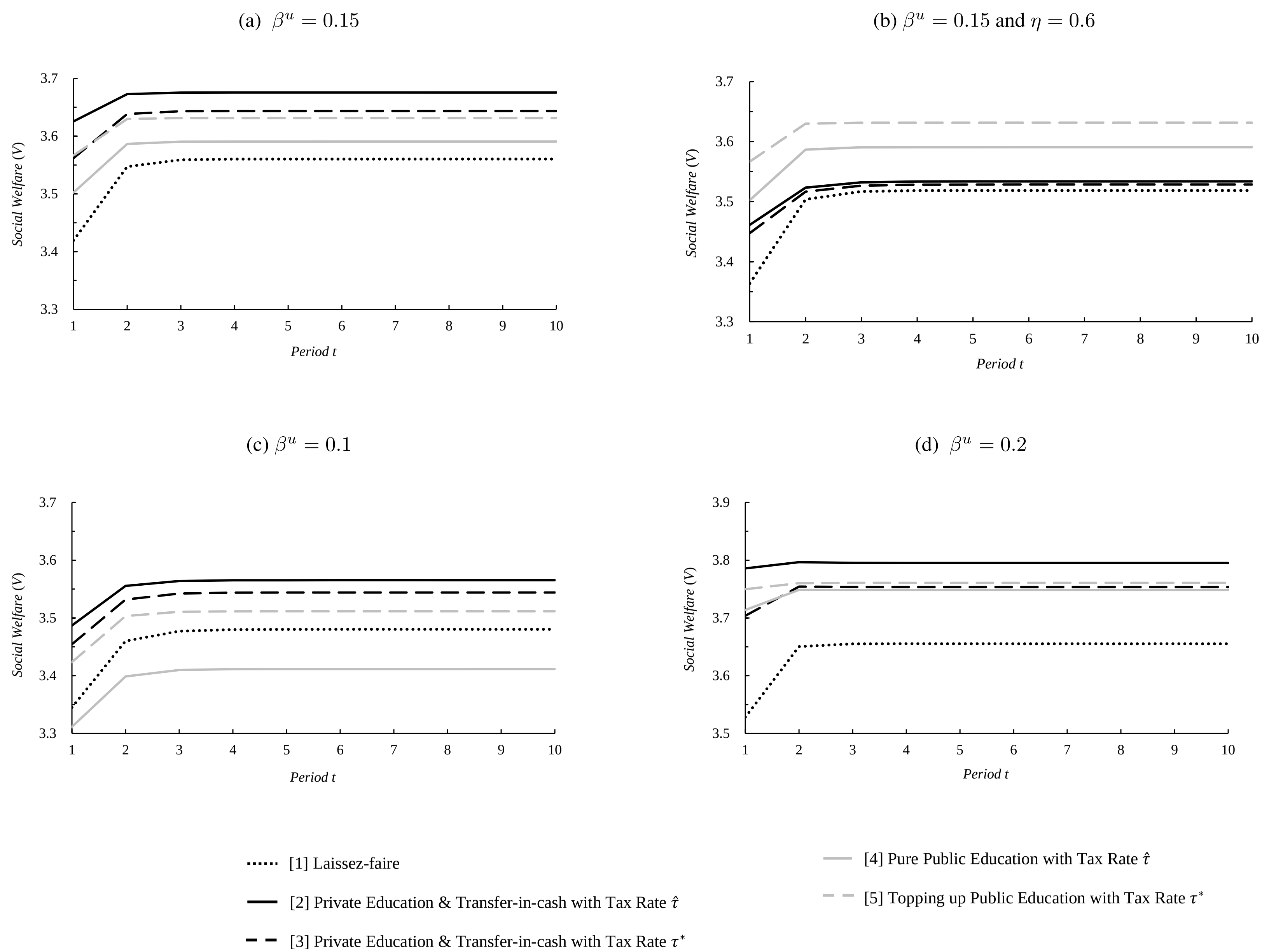
We proceed to examine the values of (M, G). Transfer in cash and transfer in kind both help to make future generations more mobile. Intergenerational mobility is at its highest in [4], which echoes Proposition 6 that the pure public schooling may provide the most equal opportunity for obtaining skilled professions. In the last column, it is intuitive that transferring wealth directly from the rich to the poor lowers earnings inequality. Nevertheless, we find a surprising result in that [4] gives rise to a greater Gini coefficient than any other system. This is largely because the public education system serves as the most effective channel for high-ability people from disadvantaged families to displace silver-spooned low-ability people in the skilled labor sector. As the increased average talent of skilled workers raises their wage rate, the between-sector pay gap widens given that the wage of unskilled workers ϕ remains constant. A downsizing of high-paid skilled workers (small θ) also helps to fuel the increase in the Gini coefficient.

We are now in a position to investigate social welfare under different policies. Our simulation analysis on social welfare have two goals. First, we aim to show the advantage of public education if some people are myopic (i.e., who care little about their children). Second, by assigning different values to β^u, we aim to illustrate how social welfare changes in response to the varying degrees of meritocracy.

Figure 2 plots the dynamic welfare consequence based on the parameter values given in Table 1 (except for β^u). Each curve exhibits an increasing trend until the steady state is achieved. In Figure 2(a), β^u is set to 0.15. It illustrates that the private school system with income transfer leads to the higher social welfare than other systems. Through the income effect, educational expenditure increases in poor families and falls in rich families, which enhances social welfare due to the law of diminishing marginal utility (e.g., Boadway and Keen 2000, Boadway and Sato 2015). This result sheds light on the practice in many countries of focusing government interference on redistributing income from the rich to the poor. However, the gray solid curve lies above the dotted curve, which means that public education is preferable to the laissez-faire outcome. As social welfare is lowest under *laissez-faire*, government intervention can always help to make people better off.

Compared with Figure 2(a), Figure 2(b) examines how the introduction of myopic people alters the simulation result. Assume that a fraction η of unskilled workers and none of skilled workers are myopic. When parents decide to make their educational investments (e.g., section 4.1), myopics will optimally choose the minimum level ($e = 1$). In that case, intergenerational social mobility

Figure 2: Welfare Consequences of Different Education Systems



amounts to

$$M_t = \frac{\lambda \ln(1 + a^H) + (1 - \eta)\beta^u \ln(\gamma\phi)}{\lambda \ln(1 + a^H) + \beta^s \ln(\gamma\delta\bar{a}_t)}. \tag{35}$$

Comparing (35) with (17) indicates that the economy becomes less mobile if myopia matters. In the presence of public schooling, the benevolent government ignores the myopic acts so that the tax and educational policies (and thus the equilibrium outcomes) are maintained as in sections 4.2 and 4.3.

Figures 2(a) and 2(b) displays different dark curves (for private education), although they share the same gray curves (for public education). In Figure 2(b), we set the proportion of myopic people in poor parents to 60% ($\eta = 0.6$). The presence of myopia makes public education more appealing: the public school system now yields a higher social welfare than any case under the private school system. The top-up system has a greater advantage and will further increase social welfare.

Figure 2(c) graphs the case in which unskilled workers perform poorly in using the educational resources ($\beta^u = 0.1$) all else being equal, while Figure 2(d) depicts the case in which they become as efficient as skilled workers ($\beta^u = 0.2$). Comparing Figure 2(a), 2(c), and 2(d) tells us which policy is socially preferable as the economy becomes increasingly meritocratic. In Figure 2(c), the most salient feature is that the pure public education system results in the lowest social welfare level (even lower than the *laissez-faire* outcome). This is because poor parents are very inefficient in using the government-funded educational resources although the implementation of public education enables their children to receive a greater amount of resources. In Figure 2(d), curve [5] lies above curve [3], implying that transfer in kind may be more desirable than transfer in cash when poor parents have no disadvantage in using educational resources.

6 Conclusion

One reason behind the populist vote in traditional European welfare states is that policies have aimed at alleviating economic inequality while neglecting the issue of social mobility. What is observed in these states is a quite stable income distribution, in great part due to proactive redistributive policies, but little concern for what has been called a broken social elevator. To illustrate, 60 years ago over 90% of parents expected their children to do better than them; today this fraction has fallen to one half.

This paper develops an overlapping-generations model analyzing the influences of educational policies on the intergenerational mobility of individuals with different innate abilities and family backgrounds. We consider that an economy's productivity and income depend on both the proportion of skilled individuals and their average talent. Our model suggests that enhancing intergenerational mobility may increase economic efficiency and equity simultaneously, which in turn increases social welfare.

In the *laissez-faire* economy where parents invest in their own children's education, the income effect suggests that the rich tend to spend more on education, leading to rich children being more likely to be rich in the future, and the poor do the opposite, giving their children fewer opportunities. Our comparative static analysis shows that, under some reasonable conditions, the proportion of skilled workers in the steady state increases with parental concern about children's educational attainment and parental income, while the average ability of skilled workers decreases with these factors. We also illustrate a Great Gatsby curve along the *dynamic* path, which implies that more equal societies tend to be more mobile over time.

We then suggest that the implementation of public education may improve social welfare, and allowing parents to make additional private educational investments on their children who attend public schools enhances social welfare further. While the public education system leads to a greater social mobility than the private education does, it may also gives rise to a higher income inequality under a certain configuration. When some poor parents are myopic, the pure public school system is strictly socially desirable. As an economy becomes more meritocratic, it is socially preferable for the government to intervene.

Appendix

Proof of Proposition 2

In the steady state, equations (19) and (20) can be rewritten as

$$\theta = \frac{\lambda \ln(1 + a^H) + \beta^u \ln(\gamma^u \phi)}{1 - \beta^s \ln(\gamma^s \delta \bar{a}) + \beta^u \ln(\gamma^u \phi)}, \tag{A.1}$$

$$\bar{a} = \lambda a^H \left[1 + \frac{(1-\lambda)\ln(1 + a^H)}{\theta}\right]. \tag{A.2}$$

Plugging (A.1) into (A.2) yields

$$\begin{aligned}
&\frac{\bar{a}}{\lambda a^H} - 1 = \frac{(1-\lambda)\ln(1+a^H)[1 - \beta^s \ln(\gamma^s \delta \bar{a}) + \beta^u \ln(\gamma^u \phi)]}{\lambda \ln(1+a^H) + \beta^u \ln(\gamma^u \phi)} \\
\Leftrightarrow\ \ &\frac{\bar{a}}{a^H} - \lambda + \left(\frac{\bar{a}}{\lambda a^H} - 1\right)\frac{\beta^u \ln(\gamma^u \phi)}{\ln(1+a^H)} = 1 - \lambda - (1-\lambda)[\beta^s \ln(\gamma^s \delta \bar{a}) - \beta^u \ln(\gamma^u \phi)] \\
\Leftrightarrow\ \ &\frac{\bar{a}}{a^H} + \frac{\bar{a}}{\lambda a^H}\frac{\beta^u \ln(\gamma^u \phi)}{\ln(1+a^H)} + (1-\lambda)[\beta^s \ln(\gamma^s \delta \bar{a})] = 1 + (1-\lambda)\beta^u \ln(\gamma^u \phi) + \frac{\beta^u \ln(\gamma^u \phi)}{\ln(1+a^H)},
\end{aligned}$$

which can be rewritten as (21). Rearranging (A.2) obtains (22).

Proof of Proposition 3

(i) Totally differentiating (21) with respect to $\bar{a}$ and a^H and then using (14) obtains

$$\begin{aligned}
&\left[1 + \frac{\beta^u \ln(\gamma^u \phi)}{\lambda \ln(1+a^H)}\right]\frac{d\bar{a}}{a^H} + \frac{\beta^s(1-\lambda)d\bar{a}}{\bar{a}} - \left\{\frac{\bar{a}}{a^{H2}} + \frac{\beta^u \bar{a} \ln(\gamma^u \phi)\left[\frac{a^H}{1+a^H} + \ln(1+a^H)\right]}{\lambda[\ln(1+a^H)a^H]^2}\right\} da^H \\
&= -\frac{\beta^u \ln(\gamma^u \phi)}{[\ln(1+a^H)]^2}\frac{da^H}{1+a^H} \\
&\Leftrightarrow\ \ a^H \Omega d\bar{a} = \left\{\bar{a} + \frac{\beta^u \ln(\gamma^u \phi)[\bar{a}a^H + \bar{a}(1+a^H)\ln(1+a^H) - \lambda a^{H2}]}{\lambda[\ln(1+a^H)]^2(1+a^H)}\right\} da^H \\
&\Leftrightarrow\ \ \frac{d\bar{a}}{da^H} = \frac{1}{a^H \Omega}\left\{\bar{a} + \frac{[\lambda(1-\lambda)a^{H2}/\theta + \bar{a}(1+a^H)]\beta^u \ln(\gamma^u \phi)}{\lambda(1+a^H)\ln(1+a^H)}\right\} > 0.
\end{aligned} \tag{A.3}$$

where, for notational simplification, we denote $\Omega := 1 + \frac{\beta^u \ln(\gamma^u \phi)}{\lambda \ln(1+a^H)} + \frac{\beta^s(1-\lambda)a^H}{\bar{a}} > 0$. Totally differentiating (21) with respect to $\bar{a}$ and λ and using (16) obtains

$$
\begin{aligned}
&\left[1+\frac{\beta^u \ln(\gamma^u\phi)}{\lambda\ln(1+a^H)}\right]\frac{d\bar{a}}{a^H}+\frac{\beta^s(1-\lambda)d\bar{a}}{\bar{a}}-\frac{\beta^u\ln(\gamma^u\phi)\bar{a}d\lambda}{\lambda^2\ln(1+a^H)a^H}=[\beta^s\ln(\gamma^s\delta\bar{a})-\beta^u\ln(\gamma^u\phi)]d\lambda\\
&\Leftrightarrow\ \frac{d\bar{a}}{d\lambda}=\frac{1}{\Omega}\left\{\frac{\beta^u\bar{a}\ln(\gamma^u\phi)}{\lambda^2\ln(1+a^H)}+a^H\Big[\beta^s\ln(\gamma^s\delta\bar{a})-\beta^u\ln(\gamma^u\phi)\Big]\right\}>0. \qquad \text{(A.4)}
\end{aligned}
$$

(ii) Totally differentiating (21) with respect to $\bar{a}$ and α and then using (16) obtains

$$
\begin{aligned}
&\left[1+\frac{\beta^u \ln(\gamma^u\phi)}{\lambda\ln(1+a^H)}\right]\frac{d\bar{a}}{a^H}+\frac{\beta^s(1-\lambda)d\bar{a}}{\bar{a}}+\frac{\bar{a}}{a^H}\frac{\beta^u}{\lambda\ln(1+a^H)}\frac{\beta^u d\alpha}{\gamma^u(\alpha\beta^u+1)^2}\\
&+\frac{\beta^s(1-\lambda)}{\gamma^s}\frac{\beta^s d\alpha}{(\alpha\beta^s+1)^2}=\beta^u\left[1-\lambda+\frac{1}{\ln(1+a^H)}\right]\frac{\beta^u d\alpha}{\gamma^u(\alpha\beta^u+1)^2}\\
&\Leftrightarrow\ \frac{\Omega d\bar{a}}{a^H}=\left[1-\lambda+\frac{1}{\ln(1+a^H)}\right]\frac{\gamma^u d\alpha}{\alpha^2}-\frac{\bar{a}}{a^H}\frac{\gamma^u d\alpha}{\alpha^2\lambda\ln(1+a^H)}-\frac{(1-\lambda)\gamma^s d\alpha}{\alpha^2}\\
&\Leftrightarrow\ \frac{d\bar{a}}{d\alpha}=-\frac{a^H}{\Omega\alpha^2}\left[(1-\lambda)(\gamma^s-\gamma^u)+\frac{\gamma^u(\bar{a}-\lambda a^H)}{\lambda a^H\ln(1+a^H)}\right]<0. \qquad \text{(A.5)}
\end{aligned}
$$

(iii) Totally differentiating (21) with respect to $\bar{a}$ and β^s obtains

$$
\begin{aligned}
&\left[1+\frac{\beta^u \ln(\gamma^u\phi)}{\lambda\ln(1+a^H)}\right]\frac{d\bar{a}}{a^H}+\frac{\beta^s(1-\lambda)d\bar{a}}{\bar{a}}+(1-\lambda)\ln(\gamma^s\delta\bar{a})d\beta^s+\frac{\beta^s(1-\lambda)\alpha d\beta^s}{\gamma^s(\alpha\beta^s+1)^2}=0\\
&\Leftrightarrow\ \frac{\Omega d\bar{a}}{a^H}=-(1-\lambda)\left[\ln(\gamma^s\delta\bar{a})+\frac{1}{\alpha\beta^s+1}\right]d\beta^s\\
&\Leftrightarrow\ \frac{d\bar{a}}{d\beta^s}=-\frac{(1-\lambda)a^H\left[\ln(\gamma^s\delta\bar{a})+1-\gamma^s\right]}{\Omega}<0. \qquad \text{(A.6)}
\end{aligned}
$$

Totally differentiating (21) with respect to $\bar{a}$ and β^u and using (14) obtains

$$
\begin{aligned}
&\left[1+\frac{\beta^u \ln(\gamma^u\phi)}{\lambda\ln(1+a^H)}\right]\frac{d\bar{a}}{a^H}+\frac{\beta^s(1-\lambda)d\bar{a}}{\bar{a}}+\frac{\bar{a}}{\lambda a^H\ln(1+a^H)}\left[\ln(\gamma^u\phi)+\frac{\alpha\beta^u}{\gamma^u(\alpha\beta^u+1)^2}\right]d\beta^u\\
&=\left[1-\lambda+\frac{1}{\ln(1+a^H)}\right]\left[\ln(\gamma^u\phi)+\frac{\alpha\beta^u}{\gamma^u(\alpha\beta^u+1)^2}\right]d\beta^u\\
&\Leftrightarrow\ \frac{\Omega d\bar{a}}{a^H}=\left[\ln(\gamma^u\phi)+\frac{1}{\alpha\beta^u+1}\right]\left[1-\lambda+\frac{\lambda a^H-\bar{a}}{\lambda a^H\ln(1+a^H)}\right]d\beta^u\\
&\Leftrightarrow\ \frac{d\bar{a}}{d\beta^u}=-\frac{a^H(1-\lambda)(1-\theta)[\ln(\gamma^u\phi)+1-\gamma^u]}{\theta\Omega}<0. \qquad \text{(A.7)}
\end{aligned}
$$

(iv) Totally differentiating (21) with respect to $\bar{a}$ and δ and rearranging obtains

$$
\left[1+\frac{\beta^u \ln(\gamma^u\phi)}{\lambda\ln(1+a^H)}\right]\frac{d\bar{a}}{a^H}+\beta^s(1-\lambda)\left(\frac{d\bar{a}}{\bar{a}}+\frac{d\delta}{\delta}\right)=0 \quad\Leftrightarrow\quad \frac{d\bar{a}}{d\delta}=-\frac{\beta^s(1-\lambda)a^H}{\delta\Omega}<0. \qquad \text{(A.8)}
$$

Totally differentiating (21) with respect to $\bar{a}$ and ϕ and using (14) obtains

$$\begin{aligned}
&\left[1+\frac{\beta^u \ln(\gamma^u \phi)}{\lambda \ln(1+a^H)}\right]\frac{d\bar{a}}{a^H}+\frac{\beta^s(1-\lambda)d\bar{a}}{\bar{a}}+\frac{\bar{a}}{a^H}\frac{\beta^u d\phi}{\lambda \ln(1+a^H)\phi}=\left[1-\lambda+\frac{1}{\ln(1+a^H)}\right]\frac{\beta^u d\phi}{\phi}\\
\Leftrightarrow\ \ &\Omega d\bar{a}=\frac{\beta^u}{\phi}\left[(1-\lambda)a^H+\frac{\lambda a^H-\bar{a}}{\lambda \ln(1+a^H)}\right]d\phi\\
\Leftrightarrow\ \ &\frac{d\bar{a}}{d\phi}=-\frac{(1-\theta)(1-\lambda)a^H\beta^u}{\theta\phi\Omega}<0. \qquad \text{(A.9)}
\end{aligned}$$

Proof of Lemma 2

All things being equal, it follows from $\beta^u \leqslant \beta^s$ and $\theta_t \in (0,1)$ that $\beta^u \leqslant \beta^s\theta_t+\beta^u(1-\theta_t) \leqslant \beta^s$. Comparing (26) with (16) obtains

$$\begin{aligned}
e_t^s-e_t^P &= \frac{\alpha\beta^s(\delta\bar{a}_t)}{\alpha\beta^s+1}-\frac{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)][\theta_t\delta\bar{a}_t+(1-\theta_t)\phi]}{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)]+1}\\
&\geqslant \frac{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)]}{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)]+1}(1-\theta_t)(\delta\bar{a}_t-\phi)>0, \qquad \text{(A.10)}\\
e_t^P-e_t^u &= \frac{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)][\theta_t\delta\bar{a}_t+(1-\theta_t)\phi]}{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)]+1}-\frac{\alpha\beta^u\phi}{\alpha\beta^u+1}\\
&\geqslant \frac{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)]}{\alpha[\beta^s\theta_t+\beta^u(1-\theta_t)]+1}\theta_t(\delta\bar{a}_t-\phi)>0. \qquad \text{(A.11)}
\end{aligned}$$

Because $e_t^s>e_t^P>e_t^u$, we use Definition 1 to derive

$$M_t^{public}=\frac{\lambda\ln(1+a^H)+\beta^u\ln e_t^P}{\lambda\ln(1+a^H)+\beta^s\ln e_t^P}>\frac{\lambda\ln(1+a^H)+\beta^u\ln e_t^u}{\lambda\ln(1+a^H)+\beta^s\ln e_t^s}=M_t^{private}.$$

Proof of Proposition 4

Inserting (26) into (12) and then focusing on the steady state obtains

$$\theta=\lambda\ln(1+a^H)+[\beta^s\theta+\beta^u(1-\theta)]\ln\left\{\frac{\lambda\delta a^H[(1-\lambda)\ln(1+a^H)+\theta]+(1-\theta)\phi}{\{\alpha[\beta^s\theta+\beta^u(1-\theta)]\}^{-1}+1}\right\}, \qquad \text{(A.12)}$$

which can be rearranged as (27). Inserting (27) into (14) and rearranging yields the steady-state average ability of skilled labor as in (28).

Proof of Proposition 5

Given $\beta^s = \beta^u = \beta$, we simplify (25) as $\tau = \frac{\alpha\beta}{\alpha\beta+1}$ and thus rewrite (27) as $K = K(\theta; \alpha, \beta, \delta, \phi, \lambda, a^H)$, which can be specified by

$$K = \theta - \lambda \ln(1 + a^H) - \beta \ln \frac{\alpha\beta}{\alpha\beta + 1} - \beta \ln \left\{ \delta\lambda a^H[(1 - \lambda) \ln(1 + a^H) + \theta] + (1 - \theta)\phi \right\} = 0.$$

We have the following four partial derivatives:

$$\frac{\partial K}{\partial \theta} = 1 - \frac{\beta(\delta\lambda a^H - \phi)}{\delta\lambda a^H[(1 - \lambda) \ln(1 + a^H) + \theta] + (1 - \theta)\phi} = 1 - \beta \frac{\delta\lambda a^H - \phi}{(\delta\bar{a} - \phi)\theta + \phi}, \tag{A.13}$$

$$\frac{\partial K}{\partial \alpha} = -\beta \cdot \frac{\alpha\beta + 1}{\alpha\beta} \cdot \frac{\beta(\alpha\beta + 1) - \alpha\beta^2}{(\alpha\beta + 1)^2} = -\frac{\beta}{\alpha(\alpha\beta + 1)} < 0, \tag{A.14}$$

$$\frac{\partial K}{\partial \delta} = -\frac{\beta\lambda a^H[(1 - \lambda) \ln(1 + a^H) + \theta]}{\delta\lambda a^H[(1 - \lambda) \ln(1 + a^H) + \theta] + (1 - \theta)\phi} < 0, \tag{A.15}$$

$$\frac{\partial K}{\partial \phi} = -\frac{\beta(1 - \theta)}{\delta\lambda a^H[(1 - \lambda) \ln(1 + a^H) + \theta] + (1 - \theta)\phi} < 0, \tag{A.16}$$

where $\frac{\partial K}{\partial \theta} > 0$ if and only if (29) is satisfied. In that case, total differentiation of K implies

$$\frac{\partial K}{\partial \theta} d\theta + \frac{\partial K}{\partial \alpha} d\alpha = 0 \qquad \Leftrightarrow \qquad \frac{d\theta}{d\alpha} = -\frac{\partial K}{\partial \alpha} \Big/ \frac{\partial K}{\partial \theta} > 0. \tag{A.17}$$

$$\frac{\partial K}{\partial \theta} d\theta + \frac{\partial K}{\partial \delta} d\delta = 0 \qquad \Leftrightarrow \qquad \frac{d\theta}{d\delta} = -\frac{\partial K}{\partial \delta} \Big/ \frac{\partial K}{\partial \theta} > 0. \tag{A.18}$$

$$\frac{\partial K}{\partial \theta} d\theta + \frac{\partial K}{\partial \phi} d\phi = 0 \qquad \Leftrightarrow \qquad \frac{d\theta}{d\phi} = -\frac{\partial K}{\partial \phi} \Big/ \frac{\partial K}{\partial \theta} > 0. \tag{A.19}$$

Because $\bar{a}$ and θ are inversely related by equation (22), we can infer that $\bar{a}$ decreases with (α, δ, ϕ) if and only if condition (29) holds.

Proof of Lemma 3

Rich children will receive private tutoring if and only if $e_t^A = \gamma^s(1 - \tau^*)\delta\bar{a}_t - \frac{\tau_t^*[\theta_t\delta\bar{a}_t + (1-\theta_t)\phi]}{(1+\alpha\beta^s)\kappa} > 0$, which can be rearranged as (34).

Proof of Proposition 6

(i) Given $\kappa, \tau_t \in (0,1)$, it is easy to see that $e_t^s = \gamma^s[\kappa(1-\tau_t)\delta\bar{a}_t - e_t^P]$ in (32) is smaller than $e_t^s = \gamma^s\delta\bar{a}_t$ in (16). Under the top-up education system, the proportional income tax leads to an effective rise in an unskilled worker's disposable income, which means that $e_t^u > \gamma^u\phi$. It follows from equations (7) and (17) that $M_t^{private} < M_t^{top-up}$.

(ii) We next compare the pure public and the top-up education systems. Holding public expenditure constant (for any given e_t^P), it is straightforward to show $M_t^{public} > M_t^{top-up}$. Nevertheless, the equilibrium public educational spending per child is larger under the pure public school system than under the top-up system for two reasons. First, as skilled workers make additional educational investments (i.e., $e^A > 0$) under the top-up system, they will pay a lower tax to support public education ("income effect"). Second, as skilled workers have already provided their children with private education under the top-up system, the marginal benefit of further educating their children in public schools gets lowered ("substitution effect").

Under the pure public school system, social mobility in (7) can be rewritten as

$$M_t^{public} = \frac{\lambda \ln(1+a^H) + \beta^u \ln e_t^P}{\lambda \ln(1+a^H) + \beta^s \ln e_t^P}. \tag{A.20}$$

Differentiating M_t^{public} with respect to e_t^P derives

$$\frac{dM_t^{public}}{de_t^P} = -\frac{(\beta^s - \beta^u)\lambda \ln(1+a^H)}{[\lambda \ln(1+a^H) + \beta^s \ln e_t^P]^2} < 0, \tag{A.21}$$

which indicates that a larger e_t^P leads to a smaller M_t^{public}. Therefore, it remains indeterminant whether $M_t^{public} > M_t^{top-up}$ in equilibrium.

References

Abbott, Brant, Giovanni Gallipoli, Costas Meghir, and Giovanni Violante. 2019. "Education policy and intergenerational transfers in equilibrium." *Journal of Political Economy*, 127(6): 2569–2624.

Alesina, Alberto, and George-Marios Angeletos. 2005. "Fairness and redistribution." *American Economic Review*, 95(4): 960–980.

Alesina, Alberto, Stefanie Stantcheva, and Edoardo Teso. 2018. "Intergenerational mobility and preferences for redistribution." *American Economic Review*, 108(2): 521–554.

Becker, Gary S., Scott Duke Kominers, Kevin Murphy, and Jorg L. Spenkuch. 2018. "A theory of intergenerational mobility." *Journal of Political Economy*, 126(S1): S7-S25.

Becker, Gary S., and Nigel Tomes. 1979. "An equilibrium theory of the distribution of income and intergenerational mobility." *Journal of Political Economy*, 87(6): 1153–1189.

Becker, Gary S., and Nigel Tomes. 1986. "Human capital and the rise and fall of families." *Journal of Labor Economics*, 4(3, Part 2): S1–S39.

Behrman, Jere R., and Mark R. Rosenzweig. 2002. "Does increasing women's schooling raise the schooling of the next generation?" *American Economic Review*, 92(1): 323–334.

Benabou, Roland. 1996. "Equity and efficiency in human capital investment: the local connection." *Review of Economic Studies*, 63(2): 237–264.

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

Boadway, Robin and Motohiro Sato, 2015. "Optimal income taxation with risky earnings: A synthesis." *Journal of Public Economic Theory*, 17(6): 773–801.

Borghans, Lex, Bart Golsteyn, James J. Heckman, and John Eric Humphries. 2016. "What grades and achievement tests measure." *Proceedings of the National Academy of Sciences*, 113(47): 13354–13359.

Brandén, Gunnar. 2019. "Does inequality reduce mobility? The Great Gatsby Curve and its mechanisms." Working Paper Series 2019:20, IFAU - Institute for Evaluation of Labour Market and Education Policy.

Bray, Mark. 2003. *Adverse Effects of Private Supplementary Tutoring: Dimensions, Implications and Government Responses.* UNESCO. Instituto Internacional de Planeamiento de la Educacion.

Bray, Mark, S. Zhan, C. Lykins, D. Wang, and O. Kwo. 2014. "Differentiated demand for private supplementary tutoring: Patterns and implications in Hong Kong secondary education." *Economics of Education Review*, 38(1): 24–37.

Carnevale, Anthony P., Megan L. Fasules, Michael C. Quinn, and Kathryn Peltier Campbell. 2019. "Born to win, schooled to lose: Why equally talented students don't get equal chances to be all they can be?" GEW Greogetown Working Paper.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review*, 104(9): 2633–2679.

Chetty, Raj, and Nathaniel Hendren. 2018. "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects." *Quarterly Journal of Economics*, 133(3): 1107–1162.

Chetty, Raj, Nathaniel Hendren, and Lawrence F. Katz. 2016. "The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment." *American Economic Review*, 106(4): 855–902.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, Emmanuel Saez, and Nicholas Turner. 2014. "Is the United States still a land of opportunity? Recent trends in intergenerational mobility." *American Economic Review*, 104(5): 141–147.

Corak, Miles. 2013. "Income inequality, equality of opportunity, and intergenerational mobility." *Journal of Economic Perspectives*, 27(3): 79–102.

Cremer, Helmuth, Philippe De Donder, Darío Maldonado, and Pierre Pestieau. 2007. "Voting over type and generosity of a pension system when some individuals are myopic." *Journal of Public Economics*, 91(10): 2041–2061.

Cremer, Helmuth, and Pierre Pestieau. 2011. "Myopia, redistribution and pensions." *European Economic Review*, 55(2): 165–175.

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 Matthias Doepke. 2009. "To segregate or to tntegrate: Education politics and democracy." *Review of Economic Studies*, 76(2): 597–628.

Diamond, Peter. 2004. "Social security." *American Economic Review*, 94(1): 1–24.

Epple, Dennis, and Richard E. Romano. 1998. "Competition between private and public schools, vouchers, and peer-group effects." *American Economic Review*, 88(1): 33–62.

Fashola, O. S. 2001. *Building Effective After School Programs.* Thousand Oaks, CA: Corwin.

Fernandez, Raquel, and Richard Rogerson. 2003. "Equity and resources: An analysis of education finance systems." *Journal of Political Economy*, 111(4): 858–897.

Fleurbaey, Marc. 1995. "Equal opportunity or equal social outcome?" *Economics and Philosophy*, 11(1): 25–55.

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

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.

Goldrick-Rab, Sara, Robert Kelchen, Douglas N. Harris, and James Benson. 2016. "Reducing income inequality in educational attainment: Experimental evidence on the impact of financial aid on college completion." *American Journal of Sociology*, 121(6): 1762–1817.

Harsanyi, John C. 1995. "Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. " *Journal of Political Economy*, 63(4): 309–321.

Hassler, John, and José V. Rodríguez Mora. 2000. "Intelligence, social mobility, and growth." *American Economic Review*, 90(4): 888–908.

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.

Jerrim, John, and Lindsey Macmillan. 2015. "Income inequality, intergenerational mobility, and the Great Gatsby Curve: Is education the key?" *Social Forces*, 94(2): 505–533.

Kim, Sunwoong, and Ju-Ho Lee. 2010. "Private tutoring and demand for education in South Korea." *Economic Development and Cultural Change*, 58(2): 259–296.

Lindbeck, Assar, and Mats Persson. 2003. "The gains from pension reform." *Journal of Economic Literature*, 41(1): 74–112.

Lucas, Robert E. 1988. "On the mechanics of economic development." *Journal of Monetary Economics*, 22: 3–42.

Neidhöfer, Guido. 2019. "Intergenerational mobility and the rise and fall of inequality: Lessons from Latin America." *Journal of Economic Inequality*, 17: 499–520.

Owens, Ann. 2016. "Inequality in children's contexts: Income segregation of households with and without children." *American Sociological Review*, 81(3): 549–574.

Owens, Ann, Sean F. Reardon, and Christopher Jencks. 2016. "Income segregation between schools and school districts." *American Educational Research Journal*, 53(4): 1159–1197.

Papageorge, Nicholas, and Kevin Thom. forthcoming. "Genes, education and labor market outcomes: Evidence from the health and retirement study." *Journal of the European Economic Association*.

Piketty, Thomas. 1995. "Social mobility and redistributive politics." *Quarterly Journal of Economics*, 110(3): 551–584.

Rawls, John. 1971. *A Theory of Justice*. Oxford, UK: Oxford University Press.

Roemer, John E. 1998. *Theories of Distributive Justice*. MA: Harvard University Press.

Roemer, John E., and Alain Trannoy. 2016. "Equality of opportunity: Theory and measurement." *Journal of Economic Literature*, 54(4): 1288–1332.

Saez, Emmanuel, and Stefanie Stantcheva. 2016. "Generalized social marginal welfare weights for optimal tax theory." *American Economic Review*, 106(1): 24–45.

Schneider, Daniel, Orestes P. Hastings, and Joe LaBriola. 2018. "Income inequality and class divides in parental investments." *American Sociological Review*, 83(3): 475–507.

Shonkoff, Jack, and Deborah Phillips. 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. National Academy Press.

Skiba, Russell, Ada Simmons, Shana Ritter, Ashley Gibb, M. K. Rausch, Jason Cuadrado, and C. G. Chung. 2008. "Achieving equity in special education: History, status, and current challenges". *Exceptional Children*, 74(3): 264–288.

Spence, Michael. 1973. "Job market signaling." *Quarterly Journal of Economics*, 87(3):355–374.

Stiglitz, Joseph E. 1975. "The theory of 'screening,' education, and the distribution of income." *American Economic Review*, 65(3): 283–300.