Contents

Poverty Alleviation through Public Education from an Income Inequality Perspective

Nuo Yi

ABSTRACT

This paper investigates the education subsidy scheme in an infinite-horizon Aiyagari-Bewley-Huggett economy with idiosyncratic human capital investment risk under linear income taxation.

We extend the heterogenous-agent in continuous-time model to a general equilibrium case through a Mean-Field-Games system combining HJB and KF equations together, and theoretically show that the stationary human capital distribution displays a Pareto type with the tail exponent depending on education elasticity and taxation. Particularly, we prove it follows an Inverse Gamma pattern. Moreover, we investigate the asymptotical linearity of high-income households' consumption and saving and find that the idiosyncratic investment risk brings larger precautionary saving. Finally, we do calibrations to fit the top-tail of 2007 U.S. income data and find out the optimal linear tax rate. To sum up, compared with the "no discrimination" tranfer economy, an increment in education subsidy scheme would make society more equal in our education-subsidy economy, while the effect of private education elasticity on human capital distribution is considerable.

Keywords: public education subsidy income inequality optimal public education

收入不平等视角下的公共教育补贴政策

衣诺

摘要

本文研究了在无限期 Aiyagari-Bewley-Huggett 经济模型中具有异质性人力资本投资风险情况下，政府采用线性所得税作为财政来源的公共教育补贴政策。

我们结合 HJB 和 KF 方程，将连续时间-异质性代理人模型从局部市场均衡拓展到一般均衡，并从理论上证明，人力资本均衡分布呈现出帕累托型，其尾部指数是教育弹性和税收的函数。具体而言，我们能够证明它在本模型中服从逆伽玛分布。此外，我们研究了居于分布右端的高收入家庭的消费、储蓄行为的渐近线性性质，并发现引入异质性投资风险确实造成了更严重的的预防性储蓄。最后，通过拟合 2007 年美国收入数据的顶端，我们进行了模型校准，并找出是的社会福利最大化的最优线性税。综上所述，与传统的“无歧视”转移支付政策相比，在教育补贴政策下人力资本的分布更为平均，并且可以通过提高教育补贴来促进社会平等，而且私人教育弹性对人力资本的分布形态的影响相当可观。

关键词： 公共教育补贴 收入不平等 最优公共教育政策

1. Introduction

Nowadays, income distributions have been shown skewed to the right by literatures on the wealth dynamics, displaying a "fat-tail" pattern with a large accumulation of mass in a relatively small range of values. In particular, the mainstream view is that the distribution of income seems to feature fat Pareto tails or power law distribution. Since human capital is the main channel to generate income, we probed into the evolution of human capital to see how it affects social welfare and income equality.

Focusing on the classic efficiency and equity issues, human capital inherently has something to do with intergenerational immobility of socioeconomic status or the persistence of economic outcomes (Black and Devereux (2010) [5]). A substantial body of research shows that economic status is persistent across generations: children raised in high-income families earn more than children raised in low-income families. What is insteresting is that human capital is also intergenerational immobility to some extend, namely, parents with higher education levels have children with higher education levels. Thus, children's education ought to be investgated for the intergenerational human capital transmission and facts like income premium is strongly associated with partental political status or family characteristics (Li et al., 2012 [19]; Black et al., 2005 [4]; Lam and Schoeni, 1993 [25]; Shea, 2000 [30]).

Talent generated genetically and education provided by family environment are two distinct channels of human capital transmission. Based on the fact that human capital transmission inequality is mainly from hetergenous parental resources (Mulligan (1997) [26]), then government intervention may be warranted on both equity and efficiency grounds. Concern over links between parental resources and children's outcomes provides a rationale for human capital redistribution programs like public education taxation, medicaid, and the earned income tax credit, such as the joint analysis of optimal income taxes and optimal human capital policies given by Stantcheva (2017) [28].

Our purpose of this article is to discuss the poverty alleviation oriented public education scheme and investigate how private education investment affects the fat-tailed distribution of income, under balanced government budget in this paper. The scheme corresponds best to compulsory education like early childhood, primary or lower secondary education, and less relevant for higher education, where loans (both public and private) are more readily available. The government wants to know what is the best public education scheme: to subsidize low-income people to what extend? Whether the expansion of education can really reduce social inequality by providing more opportunities to those from disadvantaged backgrounds, or we can only get the opposite result?

To sum up, our model builds along the same line as the work of Achdou et al. (2020) [6] (hereafter HACT (2020)), which is a continuous-time version of the ABH heterogeneous-agent model. With a Brownian diffusion process introduced in human capital accumulation, our agents are in the face of idiosyncratic risk in capital income. We study the general equilibrium under government balanced-budget constraint during which a non-linear "poverty alleviation through education" subsidy model is scrutinized and calibrated. Then, we find out the optimal tax policy based on social welfare, and then, and analysize the incidences of optimal education subsidy scheme on the economy.

Our main contributions are as follows. i) With government budget constraint holding, we make an extension of the hetergenous-agent model in continuous-time to general equilibrium. We show that there exists a tax rate optimal social welfare. ii) Our special power-return setting of educational investment (education elasticity) leads to a non-linear consumption decision in the process of solving HJB equation, which is not mentioned before. Besides, we develop an iterative method to solve KF equation apart from finite differiential algorithm. iii) We show that our ABH model is capable of generating the empirical fact that income is heavier-tailed. We theoretically show that the stationary income distribution displays a Pareto tail and the tail exponent depends on tax and power return of private education (education elasticity). iv) We explore how the scheme of education subsidy and education elasticity may alter the the tail of a stationary income distribution. Increases in education subsidy scheme and compulsory public education make society more equal. We highlight the importance of the private education elasticity in influencing income equity, compared with the elasticity of public education.

1.1 Related Literature

This paper is related to two genres of literature, the education finance and the hetergenous-agent model in continuous-time.

The government's education financing policy is a major factor in human capital redistribution. Thus there are plenty of hetergenous models studing public education scheme. Zilcha (2003) [23] studys the parental "education inclined" altruism under constant income tax and a non-discriminatory "tax-transfer" public education policy. Krebs (2003) [15] introduces the income tax and state-dependent transfer policy (poverty alleviation type), suggesting that government-sponsored severance payments to displaced workers increase growth and welfare. Bénabou (2002) [12] studys the market distortions and effects of progressive income tax and education finance, while proposing an alternative measure of pure aggregate economic efficiency in imperfect credit and insurance markets models.

Many state-of-the-art macroeconomic studies that jointly addressing aggregate and inequality issues utilize Aiyagari-Bewlet-Huggett (ABH) model since it has become a workhorse for policy evaluations. Some of them focus on explaining the fat right tail. One way to generate a Pareto tail into the ABH model for the income distribution is the introduction of idiosyncratic risk, the so-called "investment risk" (Benhabib, Bisin, and Zhu (2011, 2015) [8] [9]), which shows that stochasitc process for returns to capital stock induces a skewed distribution of income. HACT (2020) also features a fat-tailed distribution in partial equilibrium by incorporating investment risk into the canonical ABH model. The idiosyncratic risk in our paper is human capital investment risk. With the addition of government public education balance condition, we makes an extension to general equilibrium.

However, the discrete-time version of ABH economy with idiosyncratic income risk fails to explain the high observed income concentration in developed countries like the United States. (Stachurski and Toda (2019) [33]) Thus, we use the "backward-forward Mean-Field-Games" (MFG) system to solve the continuous-time version. It was initiated by Lasry and Lions (2007) [24] which describes the limiting behavior of stochastic differential problems as the number of players tends to be infinite. In MFG system, households' consumption-investment decision and the evolution of income distribution can be summarized with two differential equations, Hamilton-Jacobi-Bellman (HJB) equation and Kolmogorov Forward (KF) equation. Any heterogeneous-agent model with a continuum of atomistic agents without aggregate shock can be written as a MFG system. For example, Benhabib, Bisin and Zhu (2016) [10] studied heterogeneous-agent equilibrium models with explicit solutions. Recently, HACT(2020) has developed useful tools for solving MFG systems that do not permit closed-form solutions. The method applies as long as the heterogeneous-agent model can be boiled down to a HJB-KF combination, with any stationary Markovian process for income like continuous diffusion, and then generates a Pareto-tailed distribution.

The paper is organized as follows. In Section 2, we portray our set-ups for a tax-education subsidy model with education elasticities. We show the general equlibrium in an infinite-horizon ABH economy with idiosyncratic risk in human capital investment. In Section 3, We explain our MFG system. In Section 4, We prove the Pareto tail of income distribution and state the asymptotical proposition of high-income household's behavior. In Section 5, we show the theoretical results about policy function of households as $h \to \infty$ and the tail of the stationary income distribution, then provide the quantitative results.

2. The Model

2.1 Households

Consider an economy with one non-perishable good that can be consumed or invested. Time $t \in [0, \infty)$ is continuous. There is a continuum of infinitely-lived households of unit mass, heterogeneous in human capital/income $h > 0$ which is exogenously determined ex ante. Households are the primary economic units indexed by $i \in [0, 1]$ in computation. Each household consists of a worker and an entrepreneur: the entrepreneur runs a privately-held family business by hiring its own human capital and accumulating outputs within this household. All uncertainty is purely idiosyncratic and, therefore, all aggregates are deterministic. The state of the economy is the distribution of income.

2.2 Technology of Education

The government adopts a non-linear "poverty alleviation" subsidized redistribution education policy based on the subsidized transfer mentioned in Krebs(2003) [15], where the redistribution of public education is skewed towards the lower income. This may correspond to a policy of school funding equalization across local communities, such as poverty subsidy policies, grants and scholarships, or more generally of subsidizing differentially the education of rich and poor children.

There are two parts in public education. Government offers a constant payment $\pi < \tau \int_{\underline{h}}^{\infty} hf(h)dh$ and the subsidised transfer payment decreasing with household's income-level.[1] We assume the subsidised transfer payment follows an education subsidy scheme in power function form where constant $\beta(\zeta)$ is determined by government's tax-transfer balance. The state-dependent public education system

$$T(h, \tau) = \pi + \beta(\zeta)h^{-\zeta}, \quad b > 0 \tag{1}$$

[1]The constant payment represents the degree of equalization of public education resources, which can be regarded as the standard compulsory education for all citizens. The regressive subsidised transfer payment in total public education expenditure reflects a policy adjustment made by the gadministration to ensure the fairness of education.

All materials in education like teacher time, classrooms, books or computers are counted into the public education T and its private counterpart $e(h)$, where the marginal returns on education are $\gamma_1, \gamma_2 > 0$ respectively, while $0 < \epsilon_1, \epsilon_2 \leq 1$ capture the education elasticities. With constant endowment value w, the technology of education reads $F_e\left(e(h), T^j\right) = w + \gamma_1 e(h)^{\epsilon_1} + \gamma_2 (T^j)^{\epsilon_2}$.

2.3 Households' Consumption Decision

Assume there is one firm only use human capital to produce output $y(h) = h$. Households have identical CRRA preferences with relative risk aversion η over utility flows from future consumption discounted by factor (interest rate) ρ. Excluding consumption, we don't consider the possibility of borrowing and investing indefinitely. The household's optimization problem reads $\max\limits_{c(h),e(h)} E_0 \int_{t_0}^{\infty} e^{-\rho t} u(c) dt$ s.t.

$$c(h) + e(h) = (1-\tau)h, e(h) > 0 \tag{2}$$

$$dh = s(h)dt + \sigma h dZ, \quad s(h) = F_e(e(h), T) - \delta h \tag{3}$$

Equation 2 is the household's period budget constraint reflecting private credit market incompleteness. [1] Equation 3 is the motion of human capital. Additive idiosyncratic shock $\theta \sim N\left(0, \sigma^2\right)$ is i.i.d. across households and time and Z represents the Brownian process. [2] The saving policy function $s(.)$ represents the net human capital generated through all kinds of education investment minus depreciation, i.e. the optimally choosen drift of human capital.

2.4 Government Budget Constraint

The equilibrium public education is solved from the government budget constraint. The general government balanced-budget constraint reads $T(\tau) = \tau \int_{\underline{h}}^{\infty} y(h) f(h) dh$, i.e.

$$\int_{\underline{h}}^{\infty} \left(\pi + \beta(\zeta) h^{-\zeta}\right) f(h) dh = \tau \int_{\underline{h}}^{\infty} h f(h) dh, \quad \beta(\zeta) = \tau \frac{\int_{\underline{h}}^{\infty} h f(h) dh - \pi}{\int_{\underline{h}}^{\infty} h^{-\zeta} f(h) dh} \tag{4}$$

Saving function reads

$$s(h) = \gamma_1 \left((1-\tau)h - c(h)\right)^{\epsilon_1} + \gamma_2 \left(\pi + \beta(\zeta) h^{-\zeta}\right)^{\epsilon_2} - \delta h + w \tag{5}$$

For the sake of comparison, we also take a state-independent tax-transfer model. Government taxes gross income $y(h)$ at given tax rate τ and makes redistribution throught public education under a equilibrium state-independently transfer scheme $T(h) \geq 0$. In this kind of model, the total tax revenue from income taxation is also the lump-sum transfers. The

[1] Private credit market incompleteness can be simply interpreted by the fact that children cannot be held responsible for the debts incurred by their parents. Both consumption $c(h)$ and private education expenditure $e(h)$ are required to come out of disposable income.

[2] The scalar σ represents the undiversified idiosyncratic investment risk due to market incompleteness.

constant public education transfer for every household exactly is equal to tax revenue, thus the equilibrium total public education expenditure solves

$$T(\tau) = \tau H(\tau) = \tau \int_{\underline{h}}^{\infty} h f(\tau, h) dh \tag{6}$$

3. The Equilibrium

Heterogeneous-agent in continuous-time models can be boiled down to a MFG system consisting of two coupled PDEs. HJB equation is for the optimal choices of a single atomistic individual who takes the evolution of the distribution. KF equation characterizes the evolution of the distribution, given current optimal choices of individuals. Household's value function reads $v(h) = \max\limits_{c(h)} E_{t_0} \int_{t_0}^{\infty} e^{-\rho(t-t_0)} \frac{c(h)^{1-\eta}}{1-\eta} dt$. With borrowing limit $\underline{h}$, The domain is $(\underline{h}, \infty)$. The HJB equation reads

$$\rho v(h) = \max_{c(h)} u(c(h)) + v'(h)s(h) + \frac{\sigma^2 h^2}{2} v''(h) \tag{7}$$

Households' optimal consumption decision solves

$$c(h)^{-\eta} = v'(h)\gamma_1 \epsilon_1 \big[(1-\tau)h - c(h)\big]^{\epsilon_1 - 1} \tag{8}$$

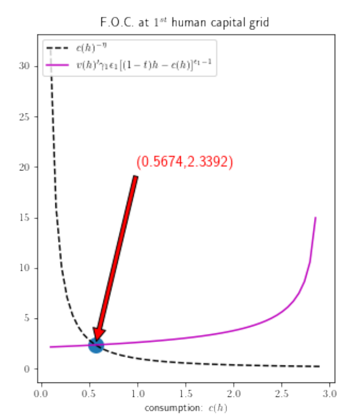

Figure 1: The Poorest

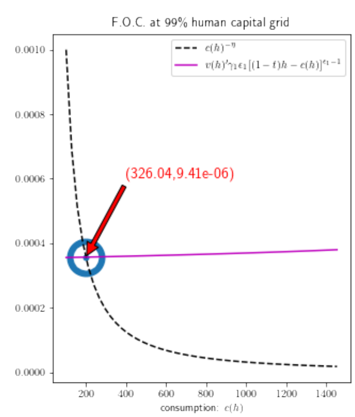

Figure 2: The Richest

Although the F.O.C. is non-linear w.r.t. $c(h)$, there exists a unique solution. [1] An example for policy functions for consumption, private education investment and saving are illustrated in figure 3 and 4.

[1]As we all know, people desire to translate their consumption from periods of high income to periods of low income to obtain more stability and predictability, instead of using all the current income to enjoy a higher standard of living, people prefer a budget that makes sense. Thus as income (human capital) increase considerably, the consumption solved through first-order-condition does not rise sharply. The phenomenon of "consumption smoothing" can also be observed in the optimal consumption decisions, such as examples in figures 1 and 2.

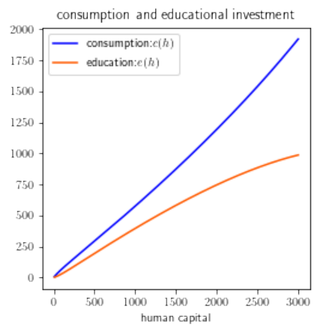

Figure 3: Consumption&Education

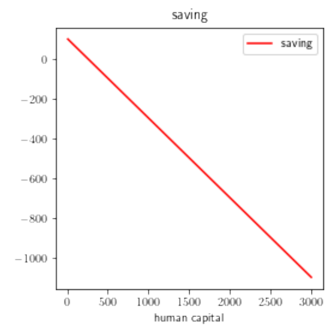

Figure 4: Saving

Let $f(.)$ denote the density function of stationary distribution which not change over time. The KF equation satisfies

$$0 = -\frac{\partial}{\partial h}\big(s(h)f(h)\big) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial h^2}\big(h^2 f(h)\big) \tag{9}$$

Proposition 3.1 *Kolmogorov forward equation written in the first-order linear differential equation reads*

$$\begin{aligned} &\frac{df(h)}{dh} + p(h)f(h) = q(h) \\ &p(h) = 2\Big[\frac{1}{h} - \frac{s(h)}{\sigma^2 h^2}\Big], \quad q(h) = \frac{2C_0}{\sigma^2 h^2} \end{aligned} \tag{10}$$

satisfing boundary conditions

$$1 = \int_{\underline{h}}^{\infty} f(h)dh \tag{11}$$

$$0 = \int_{\underline{h}}^{\infty} s(h)f(h)dh \tag{12}$$

Condition 11 is the natural property of density function. [1] Condition 12 shows that, in the stationary state, we have $\frac{d}{dh}s(h)f(h) = 0$ for all h, implying that $s(h)f(h)$ equals to a constant. With the absence of a credit market to finance human capital investment, Total saving satisfies $\int_{\underline{h}}^{\infty} s(h)f(h)dh = 0$. Additional, since first-order linear differential equation 10 is advantageous to use iterative algorithms, we develop an iterative algorithm for solving human capital density functions.

4. Household's Behavior Analysis of Income Distribution Tails

The behaviors of two tailed-groups of households still need attention. For those belonging to the bottom (left tail) of the income distributions, we refer to generically as "the low-

[1] A rigorous definition reads $1 = F(\underline{h}) + \int_{\underline{h}}^{\infty} f(h)dh$ where $F(\underline{h})$ is a potential Dirac mass at the borrowing constraint. However in our model we claim no Dirac mass, thus this condition is simplified to $1 = \int_{\underline{h}}^{\infty} f(h)dh$ or $1 = \int_{0}^{\infty} f(h)dh$. See proof in Corollary 4.1.

income", and for those that belong to the top (right tail) of the distribution, we refer to generically as "the high-income". In the empirical, this notion might be set to the bottom or top 1 percent in the gross population.

4.1 Behaviours of the Low-income Households

The saving and consumption behaviours of the low-income (at lower boundary of human capital) reflect the precautionary motive. Households suffer from idiosyncratic investment risk which cause the precautionary saving motive.

In our model, $\underline{h} = 0$ is the natural borrowing limit. To ensure the boundary constraint $h > 0$ is never violated, We need a "total employment hypothesis", which means all households in our model are in employment and $s(\underline{h}) > 0$. Thus we need a high exogenous wage assumption.

Assumption 4.1 *Exogenous wage w is sufficiently high such that the household has strong preautionary motive to ensure saving strictly larger than zero at lower human captial bound, $s(\underline{h}) > 0$.*

Then saving at lower boundary $\underline{h}$ is strictly positive. [1] The positive saving derived from "total employment hypothesis" $s(\underline{h}) > 0$ causes the lower bound of the human capital space $h = 0$ acting as a reflecting barrier of the human capital accumulation process $\{h(t)\}_{t=0}^{\infty}$. It would not be stuck at $h = 0$ and has a non-degenerating stationary distribution.

4.2 Behaviours of the High-income Households

This paper focuses on the extremely rich individuals. We will give some propositions about the high-income household's saving and consumption behavior when human capital reaches to an infinitely large value.

Lemma 4.1 *Take the derivative of HJB equation 7 w.r.t. human capital of both sides (Envelope theroem) to get a general form of Euler equation*

$$(\rho - s'(h))v'(h) = u'(c)c'(h) + v''(h)\big(s(h) + \sigma^2 h\big) + v'''(h)\frac{\sigma^2 h^2}{2} \tag{13}$$

Proposition 4.1 *The tail-proposition of the high-income household's consumption and saving behavior depends on the elasticity of private education ϵ_1. Consumption and saving have the proposition of asymptotical linearity under some parameter conditions.*

1) If $\epsilon_1 = 1$ (linear private-education return):

In the case of linear private-education return $\epsilon_1 = 1$, guess value function $v(h) = Bh^{1-\eta}$ with $B = \frac{\phi^{-\eta}}{(1-\eta)\gamma_1}$ and we get the asymptotical linearity of consumption and saving as human capital reaches infinite.

[1] Since the first-order condition of HJB equation still holds at $h = \underline{h}$, an alternative state constraint boundary condition can be $u'(c(\underline{h})) \leq v'(\underline{h})\gamma_1\epsilon_1[(1-\tau)\underline{h} - c(\underline{h})]^{\epsilon_1 - 1}$, i.e. the borrowing limit is never violated.

Euler equation reads

$$\rho = \gamma_1\big(1-\tau\big) - \delta + (a-1)\vartheta + a(a-1)\frac{\sigma^2}{2} < 0 \tag{14}$$

Consumption and saving satisfy

$$\frac{s(h)}{h} \sim \vartheta = -\frac{\rho+\delta}{\eta} + \frac{\gamma_1\big(1-\tau\big)}{\eta} - (1-\eta)\frac{\sigma^2}{2} \tag{15}$$

$$\frac{c(h)}{h} \sim \phi = \frac{(1-\tau)(\eta-1)}{\eta} + \frac{(1-\eta)\frac{\sigma^2}{2}-\delta}{\gamma_1} + \frac{\rho+\delta}{\gamma_1\eta} \tag{16}$$

2) If $2-\eta < \epsilon_1 < 1$ ***(discounted-exponent private-education return):***

In the case of discounted-exponent private-education elasticity $2-\eta < \epsilon_1 < 1$*, the value function can be termed as* $v(h) = Bh^a$*. Guess* $v(h) = Bh^{2-\eta-\epsilon_1}$ *and we get the asymptotical linearity of consumption and saving as human capital reaches infinite.*

Euler equation reads

$$\rho = -\delta + (a-1)\vartheta + a(a-1)\frac{\sigma^2}{2} \tag{17}$$

Consumption and saving satisfy

$$\frac{s(h)}{h} \sim \vartheta = -\frac{\rho+\delta}{1-a} - \frac{\sigma^2 a}{2} = \frac{\rho+\delta}{1-\eta-\epsilon_1} - \frac{2-\eta-\epsilon_1}{2}\sigma^2 \tag{18}$$

$$\frac{c(h)}{h} \sim \phi, \quad \phi^{-\eta} = (2-\eta-\epsilon_1)B\gamma_1\epsilon_1[1-\tau-\phi]^{\epsilon_1-1} \tag{19}$$

Especially when $\vartheta = -\delta$*, there exists an analytical solution*

$$\frac{s(h)}{h} \sim \vartheta = -\delta, \frac{c(h)}{h} \sim \phi = 1-\tau \tag{20}$$

3) If $\epsilon_1 < 2-\eta$ ***: there is no stationary distribution for income.*** *Guess* $v(h) = Bh^{2-\eta-\epsilon_1}$ *and it is divergent as* h *goes to infinite.*

See **Proof** in *APPENDIX 4.1.*

As $h- > \infty$, negative saving function $s(h) < 0$ is asymptotically linear on h, which induces the mean-reverting process of human capital. When $\vartheta < 0$, the consumption function $c(h)$ is asymptotically linear on h and positive. We can see that consumption is always increasing and saving always decreasing w.r.t h.

4.3 The Stationary Human Capital/Income Distribution

4.3.1 The Pareto Tail

If the absolute risk aversion remains finite at the natural borrowing limit, and relative risk aversion is bounded above for all households, i.e. $\eta = -\frac{cu''(c)}{u'(c)} < \infty, -\frac{u''(c(\tilde{0}))}{u'(c(\tilde{0}))} = \underline{R} < \infty$, then there exists a stationary equilibrium. According to MFG system, with the numerical consumption decisions solved from HJB equation, human capital distribution can be derived throught the KF equation. Now we can show that under the setting of CRRA preference, the stationary human capital distribution has a Pareto tail when $\zeta \leq 1$. The human capital distribution is smooth or the density of human capital is continuous and differentiable for all h. Moreover, an analytic expression for the tail parameter can be derived.

There are two corollaries stating that there are no Dirac points of the stationary human capital distribution.

Corollary 4.1 *Under full employment assumptions 4.1, the stationary distribution does not have a Dirac point mass at $\underline{h}$, i.e. its CDF satisfies $F(\underline{h}) = 0$, and its density is in fact finite, $f(\underline{h}) < \infty$.*

Corollary 4.2 *The support of the stationary human capital distribution is bounded above at some $h_{\max} < \infty$. Its does not have a Dirac point mass at $h_{\max}$.*

First, the condition $s(0) > 0$ must hold. If not, then it will induce the existence of unemployed people point that violates the "fully employment hypothesis". And this type of households would introduce a Dirac mass point in human capital distribution at the left tail. In our model, the stationary distribution does not have a Dirac point mass, i.e. its CDF satisfies $F(0) = 0$ or just $g(h) < \infty$ for all h. Thus $s(0)f(0)$ equals some negative constant and $s(h)f(h)$ equals to some constant for all h. Second, from proposition 4.1, we can also claim that the support of the stationary human capital distribution is bounded above at some $h_{\max} < \infty$. Thus, it does not have a mass point at both $h = 0$ and $h_{\max}$.

Definition 4.1 *If H is a random variable with a Pareto Type I distribution*[1]*, then the probability that H is greater than some certain number h. The survival function is*

$$\bar{F}(h) = \Pr(H > h) = \begin{cases} \left(\frac{h_m}{h}\right)^{\alpha} & h \geq h_{\mathrm{m}} \\ 1 & h < h_{\mathrm{m}} \end{cases} \tag{21}$$

Below are conditions for parameters in order to get stationary human capital distribution.

Assumption 4.2 *We assume that*

$$2(\rho + \delta) > 2\gamma_2(1 - \tau) + \sigma^2\eta(\eta - 1), \quad \epsilon_1 = 1 \tag{22}$$

$$2(\rho + \delta) > \sigma^2(2 - \eta - \epsilon_1)(1 - \eta - \epsilon_1), \quad 2 - \eta < \epsilon_1 < 1. \tag{23}$$

Assumption 4.2 guarantees the negativity of saving function as human capital goes to infinite, i.e. $\bar{s} = \lim_{h\to\infty} \frac{s(h)}{h} < 0$. Households suffer from Brownian idiosyncratic investment risks in human capital, which induces the precautionary saving, and push the $\bar{s}$ to be positive. See the left-hand of the equations in assumption 4.2. The higher η and σ are, the

[1] h_m (scale parameter) is the minimum possible value of H. α (shape parameter) is a positive constant named the Pareto Index (Tail Index) at around 1.5–2.5. The smaller α is, the more larger the tail-thickness is. In short, Pareto distribution describes fairly well the distribution of income exceeding a certain level in the sense that it must have a tail of order $(\frac{1}{h})^{\alpha}$.

stronger the households' precautionary savings motive is. the upper bound $2(\rho+\delta)$ confines the households' savings motive. And again, it underlines the necessary of a high-income assumption 4.1.

We claim that if households have CRRA preference, the stationary distribution has a Pareto tail with an analytic expression.

Proposition 4.2 *Under CRRA utility and "fully employment hypothesis", if $2-\eta<\epsilon_1\leq 1$, there is a unique stationary income distribution which follows an asymptotic power law, that is the p.d.f. of human capital $f(h)$ exibits a Pareto tail*

$$f(h)\sim \epsilon h^{-\Theta-1},\quad h\to\infty \tag{24}$$

where the tail exponent $\Theta = 1-\frac{2\bar{s}}{\sigma^2}$ and

$$\epsilon=\frac{2}{\sigma^2}\lim_{h\to\infty}\ln(y(h)/h^{\mu}),\quad y(h)=\frac{\sigma^2h^2f(h)}{2},\quad \mu=\frac{2\bar{s}}{\sigma^2}$$

$$\begin{aligned}\bar{s}=\bar{s}_a&=-\frac{\rho+\delta}{\eta}+\frac{\gamma_1(1-\tau)}{\eta}-(1-\eta)\frac{\sigma^2}{2}\quad(\epsilon_1=1)\\ \bar{s}=\bar{s}_b&=-\frac{\rho+\delta}{\eta-(1-\epsilon_1)}-\frac{(1-\epsilon_1)\sigma^2}{2}-(1-\eta)\frac{\sigma^2}{2}\quad(2-\eta<\epsilon_1<1)\end{aligned}.$$

This result extends Benhabib, Bisin, and Zhu (2015) [9] and HACT(2020) from a partial to a general equilibrium setting. Top income inequality $1/\Theta$ is increasing in volatility σ, risk aversion η, and decreasing in the rate of time preference ρ.

Since $\bar{s}_a>\bar{s}_b$, the top income inequality is smaller if the private education elastility is smaller than 1, $1/\Theta_a>1/\Theta_b$. The tail exponent of the income distribution under $\epsilon_1=1$ is larger than that of that under $\epsilon_1<1$. The distribution tail under linear private education return is fatter than that under non-linear private education return. Moreover, linear taxation and the introduction of education subsidy scheme has no impact on the asymptotically linear part of the policy function when $\epsilon_1<1$. This feature reflects the fact that the Pareto exponent is determined by the linear component of the policy function of the rich household.

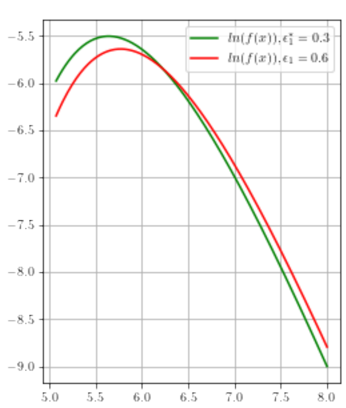

Figure 5: Check the Pareto Tail

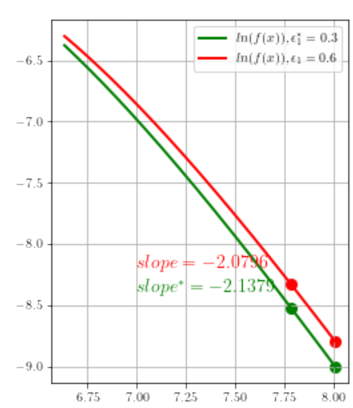

Figure 6: Compare Tails with Different ϵ_1

We use this property to examine the Pareto tail that if a variable h follows a Pareto distribution $f(h) \sim \epsilon h^{-\Theta-1}$, then its logrithm $x = \ln h$ follows an exponential distribution $f(x) \sim \epsilon e^{-\Theta x}$ and hence $\ln f(x)$ is a linear function of x where the slope equals the tail exponent Θ

$$\ln f(x)/x \sim \epsilon\Theta, \quad h \to \infty \tag{25}$$

We check out the distribution behaves asymptotically like a Pareto type by showing that the logarithm of the density $\ln f(x)$ is asymptotically linear in the logarithm of human capital $x = \ln h$. Figure 5 shows the tail exponent of stationary distribution with $\tau^* = 0.0128\%$. Figure 6 shows the slope changes w.r.t. the private education elasticity.

An incremental decrease in the elasticity of private education investment ϵ_1 leads to an incremental increase in the Pareto tail exponent, which implies a more equal distribution of human capital. Additional, public education elasticity ϵ_2 also makes some difference, the impact direction is the same as that of private education, but efficiency is much smaller.

Figures 7 and 8 shows us how the human capital distribution changes with tax rate and private education elasticity.

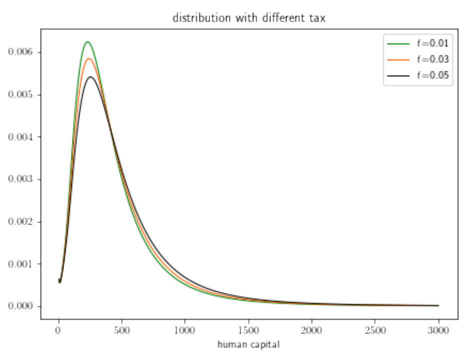

Figure 7: Change with τ

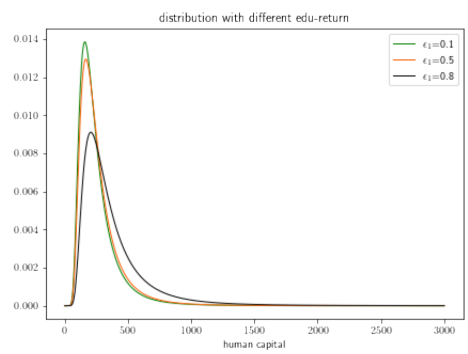

Figure 8: Change with ϵ_1

In our article, there is only idiosyncratic investment risk and the Pareto tail of the human capital distribution only depends on it. Further, if idiosyncratic labor income risk is also added, the Pareto tail will still only depend on idiosyncratic investment risk, as emphasized by Park (2020) [34] which studies optimal tax scheme that the scheme depends on the sources of income inequality. [1]

4.3.2 Inverse Gamma Pattern

We claim the human capital distribution obeies a specific Inverse Gamma pattern. Based on KF equation's first-order linear differential form 10, the p.d.f. of human capital is (where

[1] Idiosyncratic labor income instead causes the precautionary savings motive of wealth-poor households which is not mentioned in our article.

C_0, C_1 and C_2 are constants)

$$
\begin{aligned}
f(h) &= C_1 e^{-2\int_{\underline{h}}^{h} p(x)dx} + C_2 \int_{\underline{h}}^{h} x^{-2} e^{-2\int_{x}^{h} p(z)dz} dx \\
p(h) &= 2\Big[\frac{1}{h} - \frac{s(h)}{\sigma^2 h^2}\Big], \quad q(h) = \frac{2C_0}{\sigma^2 h^2}
\end{aligned}
\tag{26}
$$

Proposition 4.3 *From the boundary conditions in proposition 3.1, we can prove that* $C_2 = 0$ *and* $f(h)$ *follows an Inverse Gamma distribution*

$$
f(h) = C_1 e^{-\beta/h} h^{-\alpha-1}, \quad C_1 = \left(\int_0^{\infty} e^{-2\int_{h_0}^{h} \frac{1}{x} - \frac{s(x)}{\sigma^2 x^2} dx} dh\right)^{-1} \tag{27}
$$

See the detailed **Proof** in *APPENDIX C* .

5. Computation of Stationary Human Capital Distribution

In this section, we show the algorithm to solve MFG system equilibrium. First, the finite difference (FD) method is used to solve households' decision (HJB equations) and human capital distribution (KF equations). Second, for approximating density f, we also develop an iterative algorithm proposed in the part 4.3 to solve KF equation apart from FD method. Third, a bisection method is used to solve the balanced government budget constraint, to our criteria for determining whether the distribution of wealth has reached stationary distribution.

5.1 Finite Difference (FD) Algorithm

Based on work by Achdou and Capuzzo-Dolcetta (2012) [2], Achdou (2013) [3], and HACT (2020), We choose the FD method transforming our system of PDEs into a system of sparse matrix equations. Denote the value function $\mathbf{v}$ and distribution $\mathbf{f}$ along human capital grids $h_i, i = 1, \ldots, I$.

Upwind Scheme We choose forward or backward difference approximation according to *Upwind Scheme — using a forward difference approximation whenever the drift of the state variable is positive and the backward difference whenever it is negative.* We approximate the first-order derivative $v'(h)$ with either the forward or backward difference approximation $v'_{i,F} \approx (v_{i+1} - v_i)/\Delta h$ or $v'_{i,B} \approx (v_i - v_{i-1})/\Delta h$. The second derivative approximation is $v''(h_i) \approx (v_{i+1} - 2v_i + v_{i-1})/(\Delta h)^2$.

The end product of discretization is the following system of matrix equations

$$
\rho \mathbf{v}^{n+1} = \mathbf{u}^n(\mathbf{v}) + \mathbf{A_H}^n(\mathbf{v})\mathbf{v}^{n+1} \tag{28}
$$

$$
\mathbf{0} = \mathbf{A_K}(\mathbf{v})^{\mathrm{T}} \mathbf{f} \tag{29}
$$

The first equation reflects the discretized HJB equation 7, the second equation reflects the discretized KF equation 9. The sparse transition matrix $\mathbf{A_H}(\mathbf{v})_{I\times I}$ captures the evolution

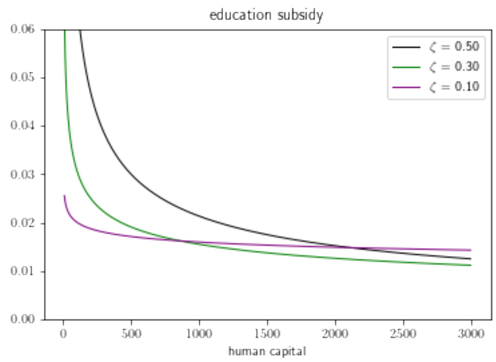

Figure 9: subsidy changes with ζ

of the idiosyncratic human capital in its discretized grids space. $\mathbf{A_K}(\mathbf{v})_{I\times I}$ in the second equation denotes the transition matrix for KF equation. [1]

There are some comparative static analysis for policy functions and value function from numerical solutions in figures 9 - 13. In the process of applying this FD method to our model, we find that the resulting density function is prone to the phenomenon of down-piling, i.e. the special mass point of the low-income type households in their setting. To test the robustness of the distribution, we develop an simplified iterative KF equation algorithm. Because our KF equation has no Poisson process, we can easily write it into the ODE-form, then apply to discretization and iteration. The density function drawn in this way has no mass point and has a fatter tail. See the detailed **Algorithm** of FD method and iterative method in *APPENDIX A*.

5.2 Calibration and Model Fit

We calibrate the parameters of our education subsidy model to replicate the income inequality in the U.S. economy. We use these parameters to calculate the quantiles of the distributions of income.

In our calibration model, we divide the parameters into two groups as in Table 1. In the first group, we calibrate the parameters from literatures. The coeffcient of relative risk aversion η is calibrated close to logarithmic utility with $\eta = 1$. We use the human capital depriciation rate (annually) from Krebs (2003) [15] as a compromise between the higher depreciation rate of physical capital and the probably lower depreciation rate of human capital. We choose the welfare-maximizing tax rate τ used in Krebs (2003). It is close to the government purchase to GDP ratio under U.S. tax (0.189) approximated by Heathcote, Store sletten, and Violante (2017) [17], a statistics regarded as an alternative of total linear taxation revenue.[2] In their definition, income includes labor earnings and capital incomes. The volatility of Brownian motion would influnce the Top 1% income shares. The higher

[1] $\mathbf{A_H}(\mathbf{v})$ and $\mathbf{A_K}(\mathbf{v})$ are derived from the HACT (2020)'s transition matrix designed for Aiyagari model with labor-income risk and investment risk in continuous time.

[2] However, nowadays, income taxation scheme has turned to progressivity type. The current progressivity of income taxation in the U.S. is 0.181 from the estimation of Heathcote et al. (2017).

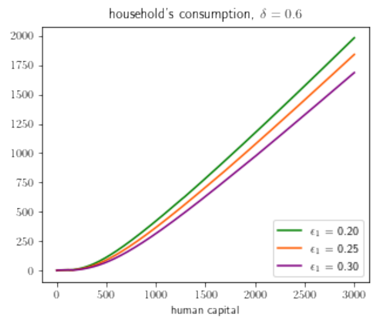

Figure 10: Consumption with Education Elasticity

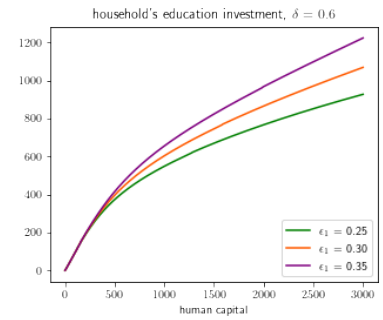

Figure 11: Education Investment with Education Elasticity

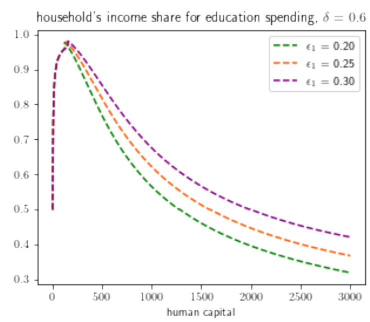

Figure 12: Education Sharing with Education Elasticity

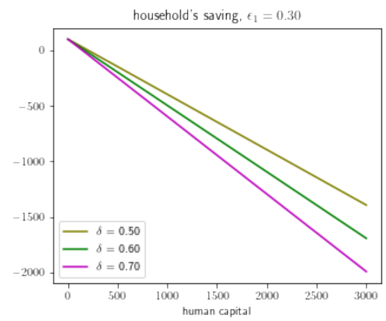

Figure 13: Saving with Depreciation

Table 1: PARAMETER TABLE

Coeffecient of Relative Risk Aversion	$\eta = 1.15$
Time Discount Factor	$\rho = 0.05$
Human Capital Depreciation Rate	$\delta = 0.06$
Linear Income Tax Rate	$\tau = 0.128$
The Volatility of Brownian Motion	$\sigma = 0.56$
Private Education Elasticity	$\epsilon_1 = 0.90$
Public Education Elasticity	$\epsilon_2 = 0.50$
Compulsory Education	$\pi = 0.10$
Education Subsidy Regressivity	$\zeta = 0.50$

volatility of Brownian motion means the higher idiosyncratic investment risk. Therefore, increasing the volatility will help to improve the income shares of Top 1 %, but the adjustment of the volatility is subject to the limitation that $\bar{s}$ must be negative.

Table 2: CALIBRATED PARAMETER VALUES

Category	Symbol	Parameter value
Preference	$u(c(h)) = \frac{c(h)^{1-\eta}}{1-\eta}$	$\eta = 1.15$
		$\rho = 0.05$
Human Capital Motion	$dh = s(h)dt + h\sigma dZ_t$	$w = 1$
		$\sigma = 0.56$
	$\theta \sim N\left(0, \sigma^2\right)$	$\gamma_1 = 0.30$
		$\gamma_2 = 0.20$
		$\delta = 0.06$
Saving Function	$s(h) = \gamma_1 e(h)^{\epsilon_1} + \gamma_2 \left(\pi + \beta(\zeta)h^{-\zeta}\right)^{\epsilon_2} - \delta h + w$	$\epsilon_1 = 0.90$
		$\epsilon_2 = 0.50$
		$\pi = 0.1$
		$\zeta = 0.50$
	$c(h) + e(h) = (1-\tau)h$	$\tau = 0.128$

Given those values, we adjust the parameters in the second group to match the targets in the U.S. data. Those targets include 8 moments of income distribution in the U.S. The calibrated parameters are summarized in Table 2, including more, like wage and education return coefficients.

We use the parameters in Tables 2 to calculate the quantiles of the distributions of income. And we compare them with 2007 U.S. data. [1] As Benhabib(2019) [11] mentions, choosing data in 2007 helps us avoid the nonstationary changes due to the Great Recession which explored in 2008. From Díaz-Giménez (2011) [18] which uses 2007 the Survey of Consumer Finances(SCF) data , income distribution of the United States has successfully ferretted out the very income-rich and wealthy. Althought the SCF sample size of 4,500 households is smaller than that of other surveys, it is particularly careful to represent the upper tail of the distributions by oversampling the rich. And our model fitting is based primarily on data from the richest people. Therefore it suits learning on fat-tailed distribution. Figure 14 displays the histogram of the income distribution.

The income distribution of our calibrated model matches the target moments reasonably well. Table 3 compares the income share moments from the data with those obtained from simulation in terms of income distribution ranking. Quantiles are values that separate fractions of the population. See the algorithm of estimating income quantiles in *APPENDIX* **??**.

Compared with the traditional tax-tranfer economy, the education subsidy scheme can generate a more equal income distribution. Generally, our model is more equal than 2007 SCF

[1] Calculated From *Facts on the distributions of earnings, income, and wealth in the united states: 2007 update.* The average income is 83,584 USD.The highest Incomes earned is around \$ 119 million. It's distributed over a large area with variance of the logs of around 0.99, coefficient of variation of around 4.32.

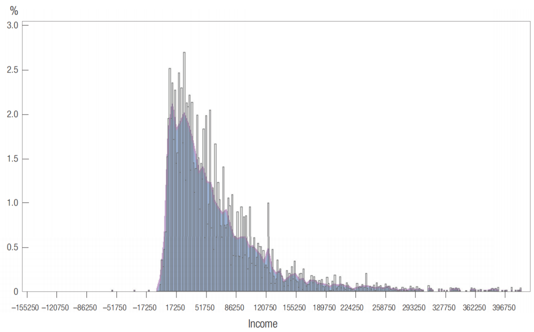

Figure 14: Income Distribution (2007 SCF, USD)

Table 3: Models Fitting based on Top1% Income Distribution

	Income Partition							
Percentile	$0-20$	$20-40$	$40-60$	$60-80$	$80-90$	$90-95$	$95-99$	$99-100$
2007 SCF data	0.028	0.067	0.113	0.183	0.138	0.102	0.159	**0.210**
Subsidy model	0.008	0.020	0.045	0.121	0.157	0.163	0.276	**0.210**
Transfer model	0.012	0.027	0.057	0.139	0.163	0.156	0.235	**0.209**

Note: Data of income share (2007 SCF) is from Díaz-Giménez, Glover, and Ríos-Rull (2011).

data, whose Gini coefficient is 0.575. The Gini coefficient of our education subsidy model is 0.264 while that of its transfer model counterpart (with same Top1% share) is 0.302. That is because a higher tax revenue in education subsidy economy will directly induces a higher education redistribution towards lower-income people, with the constant compulsory education unchanged.

Now we change some key parameters of education subsidy economy to see their incidence on human capital distribution. Table 4 shows how the income Gini coefficient changes with education elasticity and public education scheme.

The increase in education subsidy $\beta(\zeta)h^{-\zeta}$ and constant compulsory education π would, as expected, makes society more equal. Unexpectedly, it is noticable that in subsidy model, a higher tax rate does not means a more equal economy. However in tax-transfer model, an increase in linear tax rate will obviously leads to a more equal distribution. A reasonable explaination for this anomaly is that although an increase in linear taxation would reduce household's privated education investment, the decrease of private education return is much less due to the setting of the exponential educational elasticity. To be specific, private education return is concave and increasing w.r.t. private education investment. Thus for the high-income people, tax increases have relatively less negative impact on their children's private education return. From this perspective, we also explain why it is more pronounced that Gini coefficient increases w.r.t. privated education elasticity ϵ_1 than w.r.t. public education elasticity ϵ_2. As $\epsilon_1 \to 1$, the effect will be close to linear education investment return economy case. Maybe in terms of social equity, we should pay more attention to the rate of return on private education than to the contribution of public education.

On the other hand, policies are most often evaluated according to some social welfare criterion $\max_{\tau,\zeta} W(\tau) = \max \int_{\underline{h}}^{\infty} v(h)f(h,\tau)dh$.

Table 4: Gini of Subsidy Model

	Private Education Elasticity		
ϵ_1	0.7	0.8	0.9
Gini	0.077	0.202	0.264
	Public Education Elasticity		
ϵ_2	0.7	0.8	0.9
Gini	0.269	0.271	0.273
	Education Subsidy Regressivity		
ζ	0.5	1	2
Gini	0.244	0.253	0.261
	Compulsory Education		
π	0	0.1	0.5
Gini	0.269	0.264	0.252

Figure 15: $\epsilon_1 = 0.9, \sigma = 0.56$

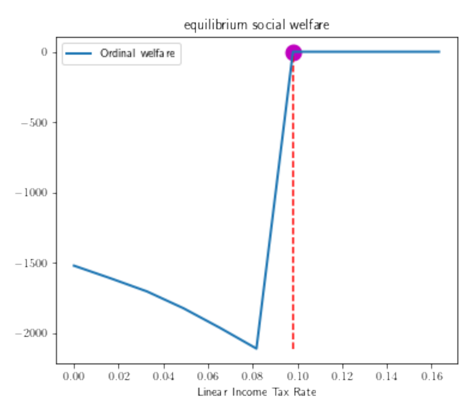

Table 5: Aggregate Analysis in Subsidy Model

	$\sigma = .56, \epsilon_1 = .7$	$\sigma = .56, \epsilon_1 = .9$	$\sigma = 0.70, \epsilon_1 = .7$	$\sigma = .70, \epsilon_1 = .9$
H	112.03(−37.49%)	179.22	54.10(−69.81%)	161.37(−9.95%)
W	−4304.01	−2577.25	−6609.95	−5842.91
S	3.081(−69.59%)	10.133	3.329(−67.14%)	15.045(48.47%)
B	0.8683(41.42%)	0.6140	0.9786(59.38%)	0.873(42.18%)

By bringing the public education scheme under the equilibrium state of each linear income tax rate into the solutions of MFG system again, we can draw the social welfare function, and then select the optimal tax for social welfare. And under the calibrated idiosyncratic risk $\sigma = 0.56$ as in figure 15, the optimal tax rate is very close to the optimal rate 0.128 in Krebs (2003) [15].

Aggregate analysis is finished on aggregate human capital / social output $H = \int_{\underline{h}}^{\infty} hf(h)dh$, aggregate saving $S = \int_{\underline{h}}^{\infty} s(h)f(h)dh$, aggregated privated edcuation $E = \int_{\underline{h}}^{\infty} e(h)f(h)dh$ and aggregated public subsidy $B = \int_{\underline{h}}^{\infty} \beta(\zeta)h^{-\zeta}f(h)dh$. Table 5 shows the aggregate analysis.[1] There are also income effects. The effect of idiosyncratic human capital investment risk σ on saving, welfare, investment, and total human capital in the education subsidy economy is similar to the effect in traditional constant transfer economy. A higher idiosyncratic risk would cause increase in precautionary savings and less private investment in education. Although the government needs to provide more education to compensate for the lack of private education, it cannot reverse the decline in overall output. The increasing elasticity of private education clearly has an incentive effect on education investment, inducing a better output. There are also substitution effects, because aggregate savings also goes up. But we can't tell how big or small these two effects are.

6. Conclusion

The U.S. income distributions have been shown skewed to the right, displaying a "fat-tail" pattern. To analyse how the public education subsidy scheme affects the income distribution of society through the evolution of human capital, we extend the hetergenous-agent in continuous-time model the using MFG systerm to a general equilibrium case. An idiosyncratic investment risk is incorporated in human capital which plays an important role in replicating the fat tail of the income distribution and induces more serious precautionary savings and government subsidies.

The calibrated model fits well with the Top 1% group fraction in 2007 U.S. income data. Compared with the traditional tax-tranfer economy, we introduce an exponential educational elasticity to our education subsidy model that can generate a more equal human capital / income distribution. The increase in "poverty alleviation by human capital" kind of education subsidy and compulsory government education would, as expected, makes society more equal.

[1] There are the gradients relative to the reference parameter $\sigma = .56, \epsilon_1 = .9$ in parentheses. Value function only reflects ordinal order, not specific utility.

References

[1] J. Aitchison and J. A. C. Brown, 1957. The Lognormal Distribution. Cambridge University Press, 1957, Pp. xviii, 176.

[2] Yves Achdou., Camilli, F., and Italo, C. D. (2012). Mean field games: Convergence of a finite difference method. Ithaca: Cornell University Library, arXiv.org.

[3] Yves Achdou. 2013. Finite difference methods for mean field games. Springer Berlin Heidelberg.

[4] Black, Sandra E., Devereux, Paul J., Salvanes, Kjell, 2005. Why the apple doesn't fall far: understanding intergenerational transmission of human capital. American Economic Review 95 (1), 437–449.

[5] Black, Sandra E., Devereux, Paul J., 2010. Recent Developments in Intergenerational Mobility. NBER Working Paper, 15889.

[6] Yves Achdou, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions and Benjamin Moll, Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach (August 2017). NBER Working Paper No. w23732.

[7] Bénabou, Roland. "Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency?" Econometrica, vol. 70, no. 2, 2002, pp. 481–517.

[8] Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2011. "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents." Econometrica 79 (1): 123–57.

[9] Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2015. "The Wealth Distribution in Bewley Economies with Capital Income Risk," Journal of Economic Theory, 159(PA), 489–515.

[10] Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2016. "The Distribution Of Wealth In The Blanchard-Yaari Model," Macroeconomic Dynamics, 20(02), 466–481.

[11] Benhabib, Jess, Alberto Bisin, and Mi Luo. 2019. "Wealth Distribution and Social Mobility in the US: A Quantitative Approach." American Economic Review, 109 (5): 1623-47.

[12] Bénabou, R. 2002. Tax and Education Policy in a Heterogeneous-Agent Economy: What Levels of Redistribution Maximize Growth and Efficiency? Econometrica, 70(2), 481-517. Retrieved July 25, 2021.

[13] Borjas, G. 1992. Ethnic Capital and Intergenerational Mobility. The Quarterly Journal of Economics, 107(1), 123-150.

[14] Christian Kleiber, and Samuel Kotz. 2003. Statistical Size Distributions in Economics and Actuarial Sciences. Frontmatter[J]. 10.1002/0471457175:i-xi.

[15] Tom Krebs. 2003. Human Capital Risk and Economic Growth, The Quarterly Journal of Economics, Volume 118, Issue 2, May, Pages 709–744.

[16] Huggett, Mark, Gustavo Ventura and Amir Yaron, 2006. "Human capital and earnings distribution dynamics," Journal of Monetary Economics, Elsevier, vol. 53(2), pages 265-290, March.

[17] Heathcote, J., K. Storesletten, and G.L. Violante. 2017. ”Optimal Tax Progressivity: An Analytical Framework,” Quarterly Journal of Economics 132(4), pp 1693–1754.

[18] Javier Diaz-Giménez, Andy Glover, José-Victor Ríos-Rull. 2011. Facts on the distributions of earnings, income, and wealth in the united states: 2007 update. Quarterly Review.

[19] Li, Hongbin, Meng, Lingsheng, Shi, Xinzheng, Binzhen, Wu., 2012. Does attending elite colleges pay in China. J. Comp. Econ. 40, 78–88.

[20] Park, Yena. 2018. Constrained Efficiency in a Human Capital Model. American Economic Journal: Macroeconomics. 10. 179-214. 10.1257/mac.20160405.

[21] Yongsung Chang and Yena Park, 2021. ”Optimal Taxation with Private Insurance,” The Review of Economic Studies, rdab003.

[22] Galo Nuno and Benjamin Moll, 2018. ”Social Optima in Economies with Heterogeneous Agents,” Review of Economic Dynamics, Elsevier for the Society for Economic Dynamics, vol. 28, pages 150-180, April.

[23] Itzhak, Zilch, 2003. Intergenerational transfers, production and income distribution. Journal of Public Economics.

[24] Lasry, J.-M., and P.-L. Lions (2007): “Mean field games,” Japanese Journal of Mathematics, 2, 229–260.

[25] Lam, David, Schoeni, Robert F., 1993. Effects of family background on earnings and returns to schooling: evidence from Brazil. Journal of Political Economy 101 (4), 710–740.

[26] Mulligan, C., 1997. Parental Priorities and Economic Inequality, University of Chicago Press, Chicago.

[27] Schiller, Kathryn S. American Journal of Sociology, vol. 100, no. 2, 1994, pp. 573–575.

[28] Stantcheva S . 2017. Optimal Taxation and Human Capital Policies over the Life Cycle[J]. Journal of Political Economy, 125(6).

[29] Saez, Emmanuel. 2001. “Using Elasticities to Derive Optimal Income Tax Rates.” The Review of Economic Studies, vol. 68, no. 1, pp. 205–229.

[30] Shea, John, 2000. Does parents’ money matter? Journal of Public Economics 77 (2), 155–184.

[31] Shavit, Yossi, Richard Arum,and Adam Gamoran.2007.Stratification in Higher Education.Palo Alto, CA:Stanford University Press.

[32] Dominik Sachs, Aleh Tsyvinski and Nicolas Werquin, 2020. ”Nonlinear Tax Incidence and Optimal Taxation in General Equilibrium,” Econometrica, Econometric Society, vol. 88(2), pages 469-493, March.

[33] Stachurski, John & Toda, Alexis Akira, 2019. ”An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity,” Journal of Economic Theory, Elsevier, vol. 182(C), pages 1-24.

[34] Chang Y , Yena P . Optimal Taxation with Private Insurance[J]. The Review of Economic Studies, 2021.

[35] Schmitt-Grohé, Stephanie and Uribe, Martín, 2001. ”Solving Dynamic General Equilibrium Models Using a Second-Order Approximation to the Policy Function,” CEPR Discussion Papers 2963, C.E.P.R. Discussion Papers.

[36] Jonathan Heathcote and Hitoshi Tsujiyama, 2021. ”Practical Optimal Income Taxation,” Staff Report 626, Federal Reserve Bank of Minneapolis.

Appendix

A. Numerical Solution of the Stationary Human Capital Distribution

A.1 FD Method Solving HJB Equation

We use an implicit FD method supporting arbitrarily step size Δ to approximate the value function $v(.)$ at I discrete points in the human capital space dimension. Denote value function by the short-hand notation $v_i \equiv v(h_i)$ and the distance between the two grid points by Δh.

A.1.1 Discretization Matrix System

According to upwind scheme, the discretization HJB equation reads $\frac{v_i^{n+1}-v_i^n}{\Delta} + \rho v_i^{n+1} = \frac{(c_i^n)^{1-\eta}}{1-\eta} + \frac{v_{i+1}^{n+1}-v_i^{n+1}}{\Delta h}(s_{i,F}^n)^+ + \frac{v_i^{n+1}-v_{i-1}^{n+1}}{\Delta h}(s_{i,B}^n)^- + \frac{\sigma^2 h_i^2 (v_{i+1}^{n+1}-2v_i^{n+1}+v_{i-1}^{n+1})}{2(\Delta h)^2}$. Collect terms with the same subscripts in the right-hand side and we end up with the discretized HJB equation

$$\frac{v_i^{n+1}-v_i^n}{\Delta} + \rho v_i^{n+1} = \frac{(c_i^n)^{1-\eta}}{1-\eta} + v_{i-1}^{n+1} x_i^n + v_i^{n+1} y_i^n + v_{i+1}^{n+1} z_i^n \tag{30}$$

$$\begin{cases} x_i^n = -\dfrac{(s_{i,B}^n)^-}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_i^2}{(\Delta h)^2} \\ y_i^n = -\dfrac{(s_{i,F}^n)^+}{\Delta h} + \dfrac{(s_{i,B}^n)^-}{\Delta h} - \sigma^2 \dfrac{h_i^2}{(\Delta h)^2} \\ z_i^n = \dfrac{(s_{i,F}^n)^+}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_i^2}{(\Delta h)^2} \end{cases} \tag{31}$$

Special Treatment for the Lower Boundary of Human Capital If we stand at the first income grid point h_1, it would certainly go forward and has no choice go backward. $(v_2^{n+1} - v_1^{n+1})(s_{1,F}^n)^+/\Delta h$ means that with savings positive, the income would increase up to the second grid h_2. And as in state boundary condition, we have $s(\underline{h}) > 0$ so that $(s_{1,B}^n)^- = 0$. Then we can substitute all $\left(v_{i,B}^n\right)'$ by $\left(v_{i,F}^n\right)'$ here for ease of calculation $\frac{v_1^{n+1}-v_1^n}{\Delta} + \rho v_1^{n+1} = u(c_1^n) + \frac{v_2^{n+1}-v_1^{n+1}}{\Delta h}\left((s_{1,F}^n)^+ + (s_{1,B}^n)^-\right) + \frac{\sigma^2}{2} h_1^2 \frac{v_3^{n+1}-2v_2^{n+1}+v_1^{n+1}}{(\Delta h)^2}$.

If there were no assets, there would be no risk volatility. Since $h_1 = 0$, the volatility term in y_1^n, z_1^n and w_1 can be drop out. Collect terms in $\frac{v_1^{n+1}-v_1^n}{\Delta} + \rho v_1^{n+1} = \frac{(c_1^n)^{1-\eta}}{1-\eta} + x_1 v_0^{n+1} + y_1^n v_1^{n+1} + z_1^n v_2^{n+1} + w_1 v_3^{n+1}$, i.e.

$$x_1 = 0, y_1^n = -\frac{(s_{1,F}^n)^+}{\Delta h}, z_1^n = \frac{(s_{1,F}^n)^+}{\Delta h}, w_1 = 0 \tag{32}$$

Special Treatment for the Upper Boundary of Human Capital For large h which referring to the people with highest income, we have $v(h) = \tilde{v}_0 + \tilde{v}_1 h^{1-\eta}$ $\quad(\epsilon_1 = 1)$ or $v(h) = \tilde{v}_0 + \tilde{v}_1 h^{2-\eta-\epsilon_1}$ $\quad(\epsilon_1 < 1)$ for unknown constants $\tilde{v}_0$ and $\tilde{v}_1$. (It is consistent with the asymptotical proposition of high-income household's behaviours stated in the proof of pareto-tailed income distribution.)

In numerical calculation, we have $h_{\max} = h_I$. Hence, we impose the following boundary condition $v''(h_I) = (1-\eta-\epsilon_1)\frac{v'(h_I)}{h_I}$. From policy function and boundary condition, we have

$$\frac{\sigma^2}{2} h_I^2 v''(h_I) = \xi v'(h_I), \xi = (1-\eta-\epsilon_1)\frac{\sigma^2}{2} h_I \tag{33}$$

The state constraint is equivalent to the situation when net wealth reaches a ceiling, consumption must be no less than saving. At the upper boundary $h_{\max} = h_I$, we make use of 33 and write the approximation as

$$\begin{aligned} \frac{v_I^{n+1} - v_I^n}{\Delta} + \rho v_I^{n+1} =& u(c_I^n) + \frac{v_{I+1}^{n+1} - v_I^{n+1}}{\Delta h}(s_{I,F}^n)^+ + \frac{v_I^{n+1} - v_{I-1}^{n+1}}{\Delta h}(s_{I,B}^n)^- \\ &+ \frac{v_I^{n+1} - v_{I-1}^{n+1}}{\Delta h}\xi, \quad \xi = (1-\eta-\epsilon_1)\frac{\sigma^2}{2} h_I \end{aligned} \tag{34}$$

As in proposition 4.1, at the upper end of the state space, $s(h) < 0$ for $h > h_{\max}$ so that $(s_{I,F}^n)^+ = 0$. Same as before, since $(s_{I,F}^n)^+ = 0$ we let $v_{I+1}^{n+1} - v_I^{n+1} = v_I^{n+1} - v_{I-1}^{n+1}$. The corresponding entries become

$$x_I^n = -\frac{(s_{I,B}^n)^-}{\Delta h} - \frac{\xi}{\Delta h}, y_I^n = \frac{(s_{I,B}^n)^-}{\Delta h} + \frac{\xi}{\Delta h}, z_I^n = 0 \tag{35}$$

Intensity Matrix Equation 30 with 31 is a system of $1 \times I$ linear equations which can be written in matrix notation with the $I \times I$ intensity matrix $\mathbf{A_H}$ as:

$$\frac{v^{n+1} - v^n}{\Delta} + \rho v^{n+1} = u^n + \mathbf{A_H}^n v^{n+1} \tag{36}$$

where

$$\mathbf{A_H}^n = \begin{bmatrix} y_1 & z_1 & w_1 & 0 & \cdots & 0 \\ x_2 & y_2 & z_2 & 0 & \ddots & \vdots \\ 0 & x_3 & y_3 & z_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 & x_I & y_I \end{bmatrix}_{I \times I}, \quad u^n = \begin{bmatrix} u\left(c_1^n\right) \\ \vdots \\ \vdots \\ \vdots \\ u\left(c_I^n\right) \end{bmatrix} \tag{37}$$

Let $\mathbf{B}^n = \left(\frac{1}{\Delta} + \rho\right)\mathbf{I} - \mathbf{A_H}^n, \quad \mathbf{b}^n = \mathbf{u}^n + \frac{1}{\Delta}\mathbf{v}^n$ where $\mathbf{I}$ is I-order identity matrix. We refer to $\mathbf{A_H}^n$ as transition matrix or intensity matrix. In particular the matrix $\mathbf{A_H}^n$ encodes the evolution of the stochastic process $\{h_t\}$ with Brownian motion. In particular, all rows sum to zero, diagonal elements are non-positive and off-diagoal elements are non-negative.

Let $\Delta \to \infty$, it is instructive to consider the case with an infinite updating step size $\frac{1}{\Delta} = 0$ and to write the HJB linear system as $\rho\mathbf{v}^{n+1} = \mathbf{u}^n + \mathbf{A_H}^n \mathbf{v}^{n+1}$. Now from the numerical solution to HJB equations we can solve the sequence of stationary consumption, saving and value function.

A.1.2 Algorithm to solve HJB equations

Step i): Initial Parameters. Guess the initial value of $c_{i,0}, v_{i,0}$ when the iteration number $n = 0$ for all human capital grids $h_{i,0} > 0, i = 1, 2, 3, ..., I-1$. The household makes the allocation decision in every period t to maximize its intertemporal utility. We need to consider is keeping $c(h)$ concave and increasing w.r.t human capital h. Moreover, under CRRA preference, we just require consumption $c(h)$ monotonically increasing.

Step ii): Update current value. Update the current value of consumption according to *Upwind Scheme*, $c_i^{n+1} = c_i^n - (s_{i,B}^n)^- - (s_{i,F}^n)^+, \quad (s_{i,F}^n)^+ = \max\{s_{i,F}^n, 0\}, \quad (s_{i,B}^n)^- = \min\{s_{i,B}^n, 0\}$. Compute current saving and value function.

Step iii): Solve v_i^{n+1} through matrix system.

A.2 FD Method Solving KF Equation

A.2.1 Discretization Matrix System

There is again a question when to use a forward & backward approximation for the derivative $[s_i f_i]'$ and $[h_i^2 f_i]''$. It turns out that the most convenient approximation is as follows

$$-\frac{(s_{i,F})^+ f_i - (s_{i-1,F})^+ f_{i-1}}{\Delta h} - \frac{(s_{i+1,B})^- f_{i+1} - (s_{i,B})^- f_i}{\Delta h} + \frac{\sigma^2}{2}\frac{h_{i+1}^2 f_{i+1} - 2h_i^2 f_i + h_{i-1}^2 f_{i-1}}{(\Delta h)^2} = 0 \tag{38}$$

Collecting terms, we have $f_{i-1}z_{i-1} + f_i y_i + f_{i+1}x_{i+1} = 0$, i.e.

$$\begin{cases} x_i = -\dfrac{\left(s_{i,B}^n\right)^-}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_i^2}{(\Delta h)^2} \\ y_i = -\dfrac{\left(s_{i,F}^n\right)^+}{\Delta h} + \dfrac{\left(s_{i,B}^n\right)^-}{\Delta h} - \sigma^2\dfrac{h_i^2}{(\Delta h)^2} \\ z_i = \dfrac{\left(s_{i,F}^n\right)^+}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_i^2}{(\Delta h)^2} \end{cases} \tag{39}$$

Special Treatment for the Lower Boundary of Human Capital At the lower boundary $h = h_{\min} = h_1$, with high-income exogeneity in our setting, there are no drifting in this boundary, and families flow out to the forward condition. Moreover, there is no mass point at lower boundary. And because of $h_1 = 0$, the Brownian motion at this grid turns to 0 and half of the Brownian motion of the forward grid flow in. $-\frac{(s_{1,F})^+ f_1}{\Delta h} + \frac{\sigma^2}{2}\frac{h^2 f_2}{(\Delta h)^2} = 0$. The special treatment for the lower boundary reads

$$z_0 = 0, y_1 = -\frac{(s_{1,F})^+}{\Delta h}, x_2 = \frac{\sigma^2 h_2^2}{2(\Delta h)^2} \tag{40}$$

Special Treatment for the Upper Boundary of Human Capital Upper Boundary At the upper boundary $h = h_{\max} = h_I$, When $\zeta \neq 0$, the cleanest solution is to impose an artificial reflecting barrier. To this end, consider the "intensity matrix" $\mathbf{A}^n$, move all entries corresponding to the (non-existent) grid point $I+1$ to the entry corresponding to I : $f_{I-1}z_{I-1}+f_I y_I+f_{I+1}x_{I+1}=0$. The special treatment for the upper boundary reads

$$\tilde{x}_I=-\frac{\left(s_{I,B}^n\right)^-}{\Delta h}+\frac{\sigma^2}{2}\frac{h_I^2}{(\Delta h)^2},\tilde{y}_I=\frac{\left(s_{I,B}^n\right)^-}{\Delta h}-\frac{\sigma^2}{2}\frac{h_I^2}{(\Delta h)^2},\tilde{z}_I=\frac{\sigma^2}{2}\frac{h_I^2}{(\Delta h)^2} \tag{41}$$

The interpretation is that whenever the process would leave the state space according to the discretized law of motion (if it would go to point $I+1$, it is ”reflected” back in (back down to point I).

Intensity Matrix Equation 38 with 39 is a system of $1\times I$ linear equations which can be written in matrix notation with the $I\times I$ intensity matrix $\mathbf{A}_k$ as $\mathbf{0}=\mathbf{A_K}^{\mathrm{T}}(\mathbf{v})\mathbf{f}$,

$$\mathbf{A_k}=\begin{bmatrix} y_1 & x_1 & 0 & 0 & \cdots & 0 \\ z_2 & y_2 & x_2 & 0 & \ddots & \cdots \\ 0 & z_3 & y_3 & x_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & z_{I-2} & y_{I-2} & x_I \\ 0 & \ddots & \ddots & 0 & z_{I-1} & y_I \end{bmatrix} \tag{42}$$

A.3 Iterative Method Solving KF Equation

Given the numerical solution of $c(h)$ from HJB numerical solutions. Human capital distribution $f(h)$ can be solved via KF equation. As described in the paper, we develop some iterative algorithms. to solve KF equation.

Intergrating the KF equation 9, we get the first-order linear differential equation form as in equation 10. An integrating factor $\mu(h)$ can be defined as $\mu(h)=e^{\int p(h)dh}=e^{2\int(\frac{1}{h}-\frac{\phi(h)}{\sigma^2h^2})dh}$. Multiply the first-order linear equation by $\mu(h)$, we get $\mu(h)\frac{df(h)}{dh}+p(h)\mu(h)f(h)=\mu(h)q(h)$. Integrating it, we get the general solution. Let $C_2=\frac{2C_0}{\sigma^2}$ and the general solution is $f(h)=C_1f_1(h)+C_2f_2(h)$, where two particular solutions satisfying

$$f_1(h)=\frac{1}{\mu(h)}=e^{-2\int_1^h(\frac{1}{x}-\frac{s(x)}{\sigma^2x^2})dx}$$
$$f_2(h)=\frac{1}{\mu(h)}\int\mu(h)q(h)dh=e^{\frac{a}{h}}h^{-b}\int e^{\frac{-a}{h}}h^{b-2}dh=e^{-2\int_1^h(\frac{1}{x}-\frac{s(x)}{\sigma^2x^2})dx}\int_1^h x^{-2}e^{2\int(\frac{1}{x}-\frac{s(z)}{\sigma^2z^2})dz}dx$$

It can be written as a variable upper bound integral form as in equation 26 where boundary conditions are the same as in equations 11 and 12.

A.3.1 Iterative Method (1)

Iterative Method (1) need us to derive the exact density formule. Since $C_1=f(h_0)$, The first method requires the exact calculation of the constant C_1. Let’s put the expression 5 and density 26 in the two boundary conditions.

From iterative method (1) we prove that Z and Z_h in the equation system 49 are non-holonomic and if the system is still true we need C_2 to be zero.[1] We calculate the value of constant C_1 throught the first line in equation system, $C_1 = 1/G = 1/\int_{\underline{h}}^{\infty}(h^{-2} - h_0^{-2})e^{2\int_{h_0}^{h}\frac{s(x)}{\sigma^2 x^2}dx}dh$. In discretization form with $C_2 = 0$, we have $f(h_{i+1}) = f(h_i) - 2\big[\frac{1}{h_i} - \frac{s(h_i)}{\sigma^2 h_i^2}\big]f(h_i)\Delta h, \quad i = 0, 1, 2, ..., I$.

We now obtain the whole sequence of $\{f(h)\}$. And in order to make sure that both formulas work, we use the density function we get to substitute it into the second equation to calculate the constant C_1' again. If $C_1 \sim C_1'$, then the equation system is satisfied.

$$C_1' = \int_{\underline{h}}^{\infty} hf(h)dh / \int_{\underline{h}}^{\infty} hg(h)dh = H(f(h))/G_h \tag{43}$$

A.3.2 Iterative Method (2)

Iterative method (2) make use of the property that the density function adds up to one, thus we do not need to know the constants in the ODE exactly. This method is easier to operate. Choose an arbitary constant $C < 1$ and let $f(1) = C$ (Finally, we normalize the density aggregation to 1.), still use the iteration equation s.t. $\sum_{i=1}^{I-1}\frac{f(h_i)+f(h_{i+1})}{2}dh = 1, \quad \sum_{i=1}^{I-1}\frac{s(h_i)f(h_i)+s(h_{i+1})f(h_{i+1})}{2}dh = 0$.

Note the distribution derived from iterative method (1) as f, and g from the iterative method (2). To check whether they are the same distribution, we use the sup-norm of the difference of two distributions, $||f-g|| = \sup_{h\in S}|f(h) - g(h)|, \quad S = [\underline{h}, \bar{h}]$. By numerical test, the norm is $||f - g|| = 8.5758 \times 10^{-17} \sim 0$. Now we can claim two methods derive the same distribution. Method 2 is used in the body of the paper.

B. Proof of Proposition 4.1

The general form of value function is $v(h) = Bh^a$, a and B are constants.

Recall the HJB equation 7 $\rho v(h) = \max_{c(h)} u(c(h)) + v'(h)\big[w + \gamma_1((1-\tau)h - c(h))^{\epsilon_1} + \gamma_2\left(\pi + \beta h^{-\zeta}\right)^{\epsilon_2} - \delta h\big] + \frac{\sigma^2 h^2}{2}v''(h)$, write it with the Hamiltonian function $H(v'(h)) = \max_{c(h)}\{u(c(h)) - v'(h)\gamma_1\big[(1-\tau)h - c(h)\big]^{\epsilon_1}\}$, then the the Euler equation $\rho v(h) = \max_{c(h)} H(v'(h)) + v'(h)\big[\gamma_2 w + \left(\pi + \beta h^{-\zeta}\right)^{\epsilon_2} - \delta h\big] + \frac{\sigma^2 h^2}{2}v''(h)$. The first-order condition equation 8 $c(h)^{-\eta} = v'(h)\gamma_1\epsilon_1\big[(1-\tau)h - c(h)\big]^{\epsilon_1 - 1}$.

B.1 If $\epsilon_1 = 1$

$$s(h) = w + \big(\gamma_1(1-\tau) - \delta\big)h - \gamma_1 c(h) + \gamma_2\left(\pi + \beta h^{-\zeta}\right)^{\epsilon_2}$$

F.O.C. solves $u'(c) = \gamma_1 v'(h)$ where $v'(h) = \frac{1}{\gamma_1}c(h)^{-\eta}, v''(h) = -\frac{\eta}{\gamma_1}c(h)^{-\eta-1}c'(h), v'''(h) = \frac{\eta(\eta+1)}{\gamma_1}c(h)^{-\eta-2}c'(h)^2 - \frac{\eta}{\gamma_1}c(h)^{-\eta-1}c''(h)$.

[1] Computing integral value, it sloves a relatively small value 1.232e-8.

Substitute them into the Euler equation in Lemma 4.1

$$\left[\frac{\rho+\delta}{\gamma_1}-(1-\tau)\right]u'(c)=\frac{1}{\gamma_1}u''(c)c'(h)s(h)+\frac{1}{\gamma_1}\frac{\sigma^2h^2}{2}\left[u'''(c)\left(c'(h)\right)^2+u''(c)c''(h)\right]+\frac{1}{\gamma_1}\sigma^2hu''(c)c'(h) \tag{44}$$

Divide both sides of the Euler equation 44 by $c(h)u''(h)$, $\left[\frac{\rho+\delta}{\gamma_1}-(1-\tau)\right]\frac{u'(c)}{c(h)u''(h)}=\frac{1}{\gamma_1}\frac{c'(h)h}{c(h)}\frac{s(h)}{h}+\frac{\sigma^2h^2}{2\gamma_1}\left[\frac{u'''(c)(c'(h))^2}{c(h)u''(h)}\right]+\frac{\sigma^2h}{\gamma_1}\frac{c'(h)}{c(h)}$. Notice that under the "asymptotical linearity consumption" assumption $\frac{c(h)}{h}\sim\phi$, we have $\lim_{h\to\infty}c'(h)=\phi$, $\lim_{h\to\infty}c''(h)=0$ and $\lim_{h\to\infty}\frac{c'(h)h}{c(h)}=1$. The relative risk aversion (RRA) $\eta=-\frac{c(h)u''(c)}{u'(c)}$ is bounded above for all c. [1] Thus we have

$$\frac{s(h)}{h}\sim\vartheta=-\frac{\rho+\delta}{\eta}+\gamma_1\frac{(1-\tau)}{\eta}-(1-\eta)\frac{\sigma^2}{2} \tag{45}$$

If us start from the general form of Euler equation rather than soving out the exact F.O.C., we can also get the same conclustion.

Given $\epsilon_1=1$, guess $a=1-\eta$ and verify $v(h)=Bh^{1-\eta}$. Hence $v'(h)=(1-\eta)Bh^{-\eta}$ and $v''(h)=-\eta(1-\eta)Bh^{-\eta-1}$. $v(h)=Bh^{1-\eta}$.

The Euler equation reads $\rho=\gamma_1\big(1-\tau\big)-\delta+(a-1)\vartheta+a(a-1)\frac{\sigma^2}{2}$. Thus the asymptotical constant ϑ of saving to human capital solves $\vartheta=-\frac{\rho+\delta}{\eta}+\frac{\gamma_1\big(1-\tau\big)}{\eta}-(1-\eta)\frac{\sigma^2}{2}$.

Since $\frac{s(h)}{h}=\frac{w}{h}+\gamma_1((1-\tau)-\frac{c(h)}{h})+\frac{\gamma_2\big(\pi+\beta h-\varsigma\big)^{\epsilon_2}}{h}-\delta,\quad h\to\infty$, we have $\vartheta=\gamma_1(1-\tau-\phi)-\delta$. Then the asymptotical constant ϕ of consumption to human capital solves $\phi=1-\tau-\frac{\vartheta+\delta}{\gamma_1}=\frac{(1-\tau)(\eta-1)}{\eta}+\frac{(1-\eta)\frac{\sigma^2}{2}-\delta}{\gamma_1}+\frac{\rho+\delta}{\gamma_1\eta}$.

Given value function, we could get the same conclusion as in equation 45. The results of the two methods are consistent and the constant coefficient of value function solves $B=\frac{\phi^{-\eta}}{\gamma_1(1-\eta)}$.

B.2 If $0<\epsilon_1<1$

Guess $a=2-\eta-\epsilon_1$ and verify $v(h)=Bh^{2-\eta-\epsilon_1}$. If we guess $c(h)$ is asymptotically linear to h, in this linear-consumption case, we note $\frac{c(h)}{h}\sim\phi$. Similarly, guess $\frac{s(h)}{h}\sim\vartheta$.

F.O.C. $u'(c)=v'(h)\gamma_1\epsilon_1\left[(1-\tau)h-c(h)\right]^{\epsilon_1-1}$. Divide the both sides of F.O.C. 8 with $h^{-\eta}$ to homogenize it

$$\left(\frac{c(h)}{h}\right)^{-\eta}=(2-\eta-\epsilon_1)B\gamma_1\epsilon_1\left[1-\tau-\frac{c(h)}{h}\right]^{\epsilon_1-1} \tag{46}$$

Then the asymptotic-ratio constant ϕ in 19 solves $\phi^{-\eta}=(2-\eta-\epsilon_1)B\gamma_1\epsilon_1[1-\tau-\phi]^{\epsilon_1-1}$. Recalling $\frac{c(h)}{h}\sim\phi$, as $h\to\infty$ we have

$$s'(h)\sim\gamma_1\epsilon_1\big(1-\tau-\phi\big)^{\epsilon_1}h^{\epsilon_1-1}-\delta \tag{47}$$

[1]Expected change of individual marginal utility of consumption can be defined as $u''(c)c'(h)s(h)=\frac{E_t\big[du'(c)\big]}{u'(c)}$.

Since $0 < \epsilon_1 < 1$, the derivative of saving as $h \to \infty$ is simplified as $s'(h) \sim -\delta$. See lemma 4.1, we get the Euler equation $\rho v'(h) = u'(c)c'(h) + v'(h)s'(h) + v''(h)\big(s(h) + \sigma^2 h\big) + v'''(h)\frac{\sigma^2 h^2}{2}$. Then Euler equation, $\rho aBh^{a-1} = aBh^{a-1}\gamma_1\epsilon_1\big[(1-\tau)h - c(h)\big]^{\epsilon_1 - 1}\phi + aBh^{a-1}s'(h) + (a-1)aBh^{a-2}\big(s(h) + \sigma^2 h\big) + (a-2)(a-1)aBh^{a-3}\frac{\sigma^2 h^2}{2}$.

Divide both sides by aBh^{a-1}, $\rho = \gamma_1\epsilon_1\phi h^{\epsilon_1 - 1}\big[(1-\tau) - \frac{c(h)}{h}\big]^{\epsilon_1 - 1} + s'(h) + (a-1)\big(\frac{s(h)}{h} + \sigma^2\big) + (a-2)(a-1)\frac{\sigma^2}{2}$. Notice $\frac{c(h)}{h} \sim \phi$, $\frac{s(h)}{h} \sim \vartheta$ and $s'(h) \sim -\delta$, as $h \to \infty$ the Euler equation reads $\rho = -\delta + (a-1)\vartheta + a(a-1)\frac{\sigma^2}{2}$, where $\vartheta = -\frac{\rho+\delta}{1-a} - \frac{\sigma^2 a}{2}$.

Since $\frac{s(h)}{h} = \frac{w}{h} + \gamma_1\big((1-\tau) - \frac{c(h)}{h}\big)^{\epsilon_1} + \frac{\gamma_2 T^{\epsilon_2}}{h} - \delta$, as $h \to \infty$, we have $\vartheta = \gamma_1 h^{\epsilon_1 - 1}((1-\tau)h^{1-\epsilon_1} - \phi)^{\epsilon_1} - \delta$, i.e. $\vartheta = -\frac{\rho+\delta}{1-a} - \frac{\sigma^2 a}{2} = \frac{\rho+\delta}{1-\eta-\epsilon_1} - \frac{2-\eta-\epsilon_1}{2}\sigma^2$.

$\phi = 1 - \tau - \frac{((\vartheta+\delta)h - w - \gamma_2 T^{\epsilon_2})^{\frac{1}{\epsilon_1}}}{h\gamma_2^{\frac{1}{\epsilon_1}}}$. Since $s'(h) = -\delta$ and $\frac{s(h)}{h} \sim \vartheta$, if $\vartheta + \delta = \frac{\rho+\delta}{1-\eta-\epsilon_1} - \frac{2-\eta-\epsilon_1}{2}\sigma^2 + \delta = 0$, $\phi = 1 - \tau + \frac{(w + \gamma_2 T^{\epsilon_2})^{\frac{1}{\epsilon_1}}}{h\gamma_2^{\frac{1}{\epsilon_1}}} \sim 1 - \tau$.

C. Proof of Proposition 4.3

Note that

$$
\begin{aligned}
g(h) &= \exp\left(-\int_{h_0}^{h} p(x)dx\right) = (h^{-2} - h_0^{-2})\exp\left(\int_{h_0}^{h} 2s(x)/(\sigma^2 x^2)dx\right) \\
z(x) &= x^{-2}\exp\left(\int_{h_0}^{x} p(z)dz\right) = x^{-2}\exp\left(\int_{h_0}^{x} 2/z - 2s(z)/(\sigma^2 z^2)dz\right) \\
G &= \int_0^\infty g(h)dh, \quad Z = \int_0^\infty g(h)\int_1^h z(x)dxdh \\
G_h &= \int_0^\infty hg(h)dh, \quad Z_h = \int_0^\infty hg(h)\int_1^h z(x)dxdh
\end{aligned}
\tag{48}
$$

Substitute the expression of human capital density in equation 26 into the two boundary conditions in proposition 3.1 and then we get the condition system as follow

$$
\begin{pmatrix} G & Z \\ G_h & Z_h \end{pmatrix}\begin{pmatrix} C_1 \\ C_2 \end{pmatrix} = \begin{pmatrix} 1 \\ H(\tau) \end{pmatrix}
\tag{49}
$$

According to the KF equation, $C_2 = 2C_0/\sigma^2$ and $C_0 = f(0)$. From the numerial method, C_2 is calculated to be zero. The mathematical method to prove it needs the prerequisite that the constant ϑ in proposition 4.1 is negative. Then we can prove that G and G_h solves certain constants as $h -> \infty$, and Z_h and Z goes to infinite. Based on this result, we claim $C_2 \sim 0, \quad h \to \infty$.

From numerical method we can prove Z and Z_h non-holonomic. We need $C_2 = 0$ to make sure the system solvable. And we'll prove it here in a mathematical way. We have known from proposition 4.1: as the human capital value goes to infinity, the saving function $s(h)$ is linear on h, $s(h) \sim \vartheta h$.

$$
\vartheta = -\frac{\rho+\delta}{\eta} + \gamma_1\frac{(1-\tau)}{\eta} - (1-\eta)\frac{\sigma^2}{2} < 0, \quad \epsilon_1 = 1
$$

$$\vartheta = -\frac{\rho+\delta}{1-a} - \frac{\sigma^2 a}{2} = \frac{\rho+\delta}{1-\eta-\epsilon_1} - \frac{2-\eta-\epsilon_1}{2}\sigma^2 < 0, \quad 2-\eta < \epsilon_1 < 1$$

Since $\vartheta = -\frac{\rho+\delta}{\eta} + \gamma_1 \frac{(1-\tau)}{\eta} - (1-\eta)\frac{\sigma^2}{2} < 0$ (For the special case $\vartheta = -\delta$, it is clearly true), it's obvious that $\vartheta - \frac{\sigma^2}{2} = -\frac{\rho+\delta}{\eta} + \gamma_1 \frac{(1-\tau)}{\eta} - (2-\eta)\frac{\sigma^2}{2} < 0, \quad \epsilon_1 = 1$. $\vartheta - \frac{\sigma^2}{2} = -\frac{\rho+\delta}{1-\eta-\epsilon_1} + \frac{1-\eta-\epsilon_1}{2}\sigma^2 < 0, \quad 2-\eta < \epsilon_1 < 1$.

Recall KF equation written as a variable upper bound integral form as the ODE in proposition 4.3, and the notes we give in equations 48.

$g(h) = h^{-2} e^{2\int_1^h \frac{s(x)}{\sigma^2 x^2} dx} \sim h^{-2} e^{2\int_1^h \frac{\vartheta}{\sigma^2 x} dx} = h^{-2+\frac{2\vartheta}{\sigma^2}}$. $h -> \infty$. The inequality condition $\vartheta < 0$, i.e. $\kappa < 0$ guarantees the convergence of the function $g(h)$ as h approaches infinity. We can claim the existence of constants G and G_h. Denote $\kappa = \frac{2\vartheta}{\sigma^2}$ and integrate functions $g(h)$ and $hg(h)$ respectively

$$G = \int_0^\infty g(h)dh = \int_0^\infty h^{-2+\frac{2\vartheta}{\sigma^2}} dh = \frac{1}{\kappa-1} h^{\kappa-1}, \quad \kappa = -1 + \frac{2\vartheta}{\sigma^2} < -1$$

$$G_h = \int_0^\infty hg(h)dh = \int_0^\infty h^{-1+\frac{2\vartheta}{\sigma^2}} dh = \frac{1}{\kappa} h^{\kappa}, \quad \kappa = \frac{2\vartheta}{\sigma^2} < 0$$

Similarly, as $x -> \infty$, $z(x) = x^{-2} e^{2\int_1^x \frac{1}{z} - \frac{s(z)}{\sigma^2 z^2} dz} \sim x^{-\frac{2\kappa}{\sigma^2}}$, where Z and Z_h goes to infinity

$$Z = \int_0^\infty g(h) \int_1^h z(x)dxdh \sim \int_0^\infty g(h) \frac{h^{1-\frac{2\vartheta}{\sigma^2}}}{1-\frac{2\vartheta}{\sigma^2}} dh = \frac{1}{1-\frac{2\vartheta}{\sigma^2}} \int_0^\infty \frac{1}{h} dh = \ln h^{1-\kappa}, \quad 1-\kappa > 1$$

$$Z_h = \int_0^\infty hg(h) \int_1^h z(x)dxdh \sim \int_0^\infty g(h) \frac{h^{2-\frac{2\vartheta}{\sigma^2}}}{1-\frac{2\vartheta}{\sigma^2}} dh = \frac{1}{1-\frac{2\vartheta}{\sigma^2}} \int_0^\infty 1 dh = \ln h^{2-\kappa}, \quad 2-\kappa > 0$$

Recall the equation system 49, $C_2 = \frac{2C_0}{\sigma^2}$ and $C_0 = f(\underline{h})$, G and G_h solve certain constants as and Z_h and Z are both infinite value, we can claim that $C_2 \sim 0, h -> \infty$.

Proof of Inverse Gamma Distribution Kernel Gamma function is defined as $\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt$ with shape parameter α and scale parameter β. Inverse Gamma distribution is the reciprocal of a Gamma variable, the core of which is $e^{-\frac{\beta}{h}} h^{-\alpha-1}$. If x obeys Inverse Gamma distribution, the probability density function defined over the support $h > 0$ is $f(h) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\frac{\beta}{h}} h^{-\alpha-1} = Ce^{-\frac{\beta}{h}} h^{-\alpha-1}$, where integration constant is C. We can compare it with the result in proposition 4.3.

.

D. Proof of Proposition 4.2

Intergrating KF equation 9, we get $\frac{\sigma^2}{2}\frac{\partial}{\partial h}[h^2 f(h)] = s(h)f(h) + C_0$. As there is no mass point on the lower boundary, we choose $C_0 = 0$ as an implicit boundary condition. From proposition 4.1, we have $s(h) = \tilde{s} + \bar{s}h$ depending on the value of private education elasticity, where $\bar{s}_a = -\frac{\rho+\delta}{\eta} + \frac{\gamma_1 (1-\tau)}{\eta} - (1-\eta)\frac{\sigma^2}{2} \quad (\epsilon_1 = 1)$, $\bar{s}_b = -\frac{\rho+\delta}{\eta-(1-\epsilon_1)} - \frac{(1-\epsilon_1)\sigma^2}{2} - (1-\eta)\frac{\sigma^2}{2} \quad (2-\eta < \epsilon_1 < 1)$, Now we define $y(h) = \frac{\sigma^2 h^2 f(h)}{2}$, and rewrite KF equation as $y'(h) = \frac{2s(h)}{\sigma^2 h^2} y(h) = 2\frac{\tilde{s}+\bar{s}h}{\sigma^2 h^2} y(h) = \mu_1 \frac{y(h)}{h} + \mu_2(h)y(h)$. where $\mu_1 = \frac{2\bar{s}}{\sigma^2}, \mu_2(h) = \frac{2\tilde{s}}{\sigma^2 h^2} = 2\frac{\tilde{s}+\bar{s}h}{\sigma^2 h^2} - 2\frac{\bar{s}}{\sigma^2 h}$. Dividing KF equation by $y(h)$, $\frac{y'(h)}{y(h)} = \frac{\mu_1}{h} + \mu_2(h)$. Intergrating both side w.r.t h, we have $\ln(\frac{y(h)}{h^{\mu_1}}) = \int_{\underline{h}}^{\bar{h}} \mu_2(h)dh = \int_{\underline{h}}^{\bar{h}} \frac{2\tilde{s}}{\sigma^2 h^2} dh = \frac{-2\tilde{s}}{\sigma^2} h^{-1}|_{\underline{h}}^{\bar{h}}$. Hence there exist $\bar{\epsilon}$ s.t. $\lim_{h\to\infty} \ln(\frac{y(h)}{h^{\mu_1}}) = \bar{\epsilon}$.

Recall the definition $y(h) = \frac{\sigma^2 h^2 f(h)}{2}$, we have $\lim_{h\to\infty} \frac{f(h)}{h^{\mu_1 - 2}} = \frac{2\exp(\bar{\epsilon})}{\sigma^2}$. To sum up, we claim when $h \to \infty$, $f(h) \sim \epsilon h^{-\zeta-1}, \epsilon = \frac{2\bar{\epsilon}}{\sigma^2}, \zeta = 1 - \mu_1 = 1 - \frac{2\bar{s}}{\sigma^2}$.

Acknowledgements

My interest in economics, particularly in the sub-genre of macroeconomics, budded in my first class in college. I felt marveled when being informed that the essence of macroeconomics was to work out the matter on social felicity. It was the first time my curiosity for locating the underlying cause for correlation between the two was aroused, and that curiosity amplified as my research on the field went further.

Years-long study in Economics (Honor Program) at UIBE led me to unshackle the seemingly-complicated concept of macroeconomics. Macroeconomics, in my understanding, is to study and work out allocation of issues before redressing all manners of civil issues such as employment, guarantee, tax, and so on. A country with a bleak economy on the macroeconomic level would witness crises of all kinds ranging from debt crisis, economic crisis to humanitarian crisis that would lead to unemployment, idle resources, reduced productivity, and more. People's bliss would thus be sapped, and the vulnerable would be subject to more grilling torment. This understanding has consolidated my determination to pursue a career in public sector to form better policies on the macroeconomic level. In this way, I can contribute to the welfare of people, particularly those in dire need of help.

I would like to acknowledge Professor Shenghao Zhu, who set an example of excellence as a researcher, for his kindly support and wise advices. He provided me the first opportunity in my life to get in touch with serious theoretical research in macroeconomics. Tutored by Prof. Zhu, I finished the research on effects of public education schemes on 'fat-tail' income distribution in ABH economies with idiosyncratic human capital investment risks. By the way, I'd like to thank my groupmates Yitong Wang and Xueya Zhao who gave great help and sincere encourage in my hard days.

Although there were so many difficuties en route to finish my research, I figure that all those practices, boosted my programming and modelling capacities further. This experience made me attempt to pursue macroeconomic research from a computer-related perspective hereafter. An intense urge to pursue doctorate study began to creep up in me.

The last blessing is for myself: whether engaged in scientific research in the future or not, be a person who has ardent awe of her career, grateful heart to her benefactor and love for her motherland.