Poverty Alleviation through Public Education from an Income Inequality Perspective

Nuo Yi *

November 13, 2021

Abstract

This paper investigates the education elasiticity and education subsidy scheme in an infinite-horizon Aiyagari-Bewley Huggett(ABH) economy with idiosyncratic human capital investment risk under linear income taxation. First, through Mean-Field-Games (MFG) system combining Hamilton-Jacobi-Bellman(HJB) and Kolmogorov Forward(KF) equations, we theoretically show that the stationary human capital distribution displays a Pareto tail and the tail exponent depends on education elasiticity and taxation. Particularly, we show it follows an Inverse Gamma pattern. Second, we prove consumption and saving behaviors' asymptotical linearity propositions of the high-income households. Third, we do calibrations to fit the 2007 U.S. data based on the income tail and find out the optimal linear tax rate. Compared with the traditional tax-tranfer economy, our subsidy economy generate income distributions where increases in education subsidy scheme and compulsory public education would, as expected, make society more equal. We find that idiosyncratic investment risk would bring more precautionary savings and there are both income effects and substitution effects of educational elasticity variation.

Keywords human capital accumulation; education subsidy; education elasticity; income distribution; optimal income tax;

1 Introduction

This paper addresses education elasticity and public education subsidy under balanced government budget, investigating how household's private education investment affects income distribution. Our public education scheme is based on the subsidized payment

*Nuo Yi, B.S. in Economics, University of International Business and Economics. Email to: 201801030@uibe.edu.cn. See related work and codes on Github:https://github.com/yinuonino.

in Krebs (2003) [15]. The difference is that we take a regressive subsidy form decreasing with households' income and we propose a power-return form of educational investment (education elasticity) to diversify the investment effects in education.

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.[1] 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 unobservable, we use income as a proxy variable of human capital. [2]

To sum up, our model economy builds along the same line as the work of Achdou, Han, Lasry, Lions, and Moll (2020)[6] (hereafter HACT(2020)), which is a continuous-time version of the ABH heterogeneous agent model. We introduce a Brownian diffusion income process in human capital accumulation to feature such fat Pareto-tailed stationary distribution, where agents are in the face of idiosyncratic risk in capital income instead of in labor income. We study public education scheme equilibriums under government balanced-budget constraint during which a state-independent transfer model and a non-linear "poverty alleviation through education" subsidy model are scrutinized and calibrated. Our purpose of this article is finding out the optimal tax policy based on social welfare when human capital is regarded as a kind of risk capital, and then, investgating optimal education subsidy scheme and analysizing its incidence on the economy.

Our main contributions are as follows. (1) We show that with government budget constraint hold, there exists a tax rate optimal social welfare. We make an extension to general equilibrium and during computation, we develop an iterative method to generate human capital distribution apart from original algorithm in HACT(2020). Besides, our special power-return setting of educational investment(education elasticity) leads to a non-linear first-order condition in the process of solving HJB equation, which is not mentioned before. (2) 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 would, as expected, make society more equal. We highlight the importance of substantial returns of private education in influencing income equity, compared with than of public education. (3) We show that our non-canonical ABH model is capable of generating the empirical fact that wealth is heavier-tailed under the assumptions: i) infinitely-lived agents, ii) risk-free saving , and iii) constant discount factors. We theoretically show that the stationary income distribution displays a Pareto tail and the tail exponent depends on the tax rate and power return of private education.

[1] Díaz-Giménez et al. (2011) [18] gives a description of inequality in the United States in 2007 calculated by the 2007 United States' Survey of Consumer Finances (SCF) data. It also shows that joint distribution of wealth and earning are in the same pattern.

[2] Income is defined as the payments to the factors of production owned by the household plus transfers. It consists of all kinds of revenue before taxes, including earnings plus capital income plus transfers.

1.1 Human Capital as Risky Asset

During the past two decades, the Aiyagari-Bewlet-Huggett (ABH) model has become a workhorse for policy evaluations in the state-of-the-art macroeconomics that jointly addresses aggregate and inequality issues. However, discrete-time versions of ABH economies with single idiosyncratic labor income risk fail to explain the high observed wealth concentration in developed countries like the United States. (Stachurski and Toda (2019) [34])

Many recent studies of inequality focus on the relatively difficult task of explaining the thickness of the upper tail. One important factor is capital risk.[1] The introduction of idiosyncratic risk in capital income or the so-called "investment risk" into the ABH model to generate a Pareto tail for the wealth distribution is due to Benhabib, Bisin, and Zhu (2011, 2015).[8] [9] , which show that stochasitc process for returns to capital stock induces a skewed distribution of wealth. [2] Propositions 9-10 in HACT(2020) also feature a fat-tailed stationary wealth distribution by incorporating investment risk into the canonical ABH model. To our knowledge, however, all of the generation of a Pareto tail for the income distribution under investment risk are conducted under the setting of partial rather than general equilibrium. Our paper makes an extension to general equilibrium.

In this article, we use a system called the "backward-forward Mean-Field-Games" (MFG) system initiated by Lasry and Lions (2007)[24]. [3] In MFG system, households' consumption-investment decision and the evolution of income distribution[4] can be summarized with two differential equations: a Hamilton-Jacobi-Bellman (HJB) equation and a Kolmogorov Forward (KF) equation.[5] Any heterogeneous agent model with a continuum of atomistic agents (and without aggregate shocks) can be written as a MFG system. For example, Benhabib, Bisin and Zhu (2016) [10] have studied heterogeneous-agent equilibrium models with explicit solutions in continuous time, yielding a system of coupled HJB and KF equations. HACT(2020) also develop useful tools for solving and analyzing models that do not permit closed-form solutions. It 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 (e.g., Brownian process) or jump-diffusion processes (e.g., Poisson process), it would generate a Pareto-tailed distribution. In our article, we find the statioanry distribution of the general equilbirum through solving this couple system of HJB amd KF equations. Calibration exercises of the model obtain income distributions which match the quantile and top-percent shares

[1] There are other way of explaining. For instance, the Courtadon process (Courtadon,1982) also provides a mechanism to get the fat-tailed distribution, widely used in finance literatures. One interesting feature of the Courtadon process is that the Geometric Brownian Motion is a special nested case.

[2] Moreover, Benhabib et al. (2019) [11] further proves the capital risk fits best the cross-sectional income distribution in the United States out of three potential factors.

[3] Generally, their approach is to cast heterogeneous agent models in terms of the mathematical theory of MFG.

[4] or the joint distribution of their income and wealth

[5] More on these two equations, see Achdou, et al. (2020).

of the U.S. data.

1.2 Poverty Alleviation through Public Education

Like our setting of endogenous educational investment, there are many literatures provide analytical solutions builded on some hetergenous agent models in discrete time, with stochastic human capital accumulation processes. Borjas (1992) [13] studies ethnicity's role in intergenerational mobility as an externality from human capital accumulation. Huggett et al.(2006) [16] develop an model with endogenous human capital accumulation to match many features of the earning distribution. By contrast, there are also literatures focusing on representative household model like Grossman and Helpman (2020). [1]

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.(Haveman and Wolfe (1995), Mayer (1997)) 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 [2]. 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]; Plug, 2004; Shea, 2000[31]).

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)[27]), then government intervention may be warranted on both equity and efficiency grounds.(Benabou (1996)) 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 [3]. Stantcheva(2017)[29] gave a joint analysis of optimal income taxes and optimal human capital policies in a life cycle model.

In this article, we will discuiss the public education program given a linear income taxation. It might correspond best to compulsory education like early childhood, elementary and secondary schooling. It is less relevant for higher education, where loans (both public and private) are more readily available. Main problem is that the govern-

[1] It shows a balanced growth path exists and is characterized by an inverse relationship between the capital/labor-augmenting technological progress rates and the capital share in GDP.

[2] For example, Black et al. (2005) uses the reform of the education system in Norway to prove the high correlations between parents' and children's education. Lam and Schoeni (1993) finds the omitted family background variables are important for explaining high estimated returns to schooling in Brazil.

[3] Thus some research has been done from the perspective of income distribution like (Loury(1981)[26]).

ment wants to know what is the best public education program: to subsidize low-income people, to support highly educated people or something else? Whether the expansion of higher education reduces social inequality by providing more opportunities to those from disadvantaged backgrounds, or whether it mainly benefits those from advantaged families, thereby widening it? [1]

1.3 Related Literature

This paper is related to the literatures on distortionary income taxation and education finance. The government's education financing policy is a major factor in human capital redistribution. There are plenty of modes of educational "tax-transfer". Zilcha(2003) [23] studies the parental "education inclined" altruism on distribution under income-tax and constant-transfer public education policy. Krebs (2003) [15] introduces the income-tax and state-dependent transfer policy (subsidized payment, increasing on the human capital), suggesting that government-sponsored severance payments to displaced workers increase growth and welfare. Bénabou (2002) [12] studied the market distortions and effects of progressive income taxes and education finance in a dynamic heterogeneous-agent economy, while proposing an alternative measure of pure aggregate economic efficiency in imperfect credit and insurance markets models.

The paper is organized as follows. In Section 2, we portray our set-ups for a state-independent tax-transfer model and a tax-education subsidy model both with education elasticities. We show their 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 behaviour. 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 Set-up

2.1 Basic Model: State-independent Tax-Transfer Model

Consider an economy with one non-perishable good that can be consumed or invested. Time is continuous, indexed by $t \in [0, \infty)$. 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

[1] Shavit and Blossfeld (1994) shows the expansion of higher education increases inequality. One flaw is the multiple-countries study lacks empirical research from China.

computation. [2] Here we can assume that a couple bring up two children keeping the population unchanged. 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/human capital.

Households have identical constant relative risk aversion (CRRA) preferences $u(c_t) = \frac{c_t^{1-\eta}-1}{1-\eta}$ with relative risk aversion η over utility flows from future consumption discounted by factor (interest rate) ρ. [1] Excluding consumption, we don't consider the possibility of borrowing and investing indefinitely.

Given linear income tax rate τ, the government taxes gross income $y(h)$ at tax rate τ and makes redistribution throught public education subsidy under a equilibrium state-independently transfer scheme (public education) $T(h) \geq 0$. The equilibrium public education is solved from the government budget constraint. In order to simplify, here $T(h)$ can be regarded as a constant T for all households.

A household's lifetime utility maximization problem reads

$$\max_{c(h),e(h)} E_0 \int_{t_0}^{\infty} e^{-\rho t} u(c) dt \tag{1}$$

s.t.

$$y(h) = h \tag{2}$$

$$c(h) + e(h) = (1-\tau)y(h), e(h) > 0 \tag{3}$$

$$dh = (w + \gamma_1 e(h)^{\epsilon_1} + \gamma_2 T^{\epsilon_2} - \delta h)\, dt + h\sigma dZ \tag{4}$$

Equation 2 is the production technology which can be explained as there is one firm only use human capital to produce output $y(h)$ [2]. The resulting income $y(h)$ is then subject to taxes and transfers, resulting in a disposable income denoted $(1-\tau)y(h)$.

---

[2] A primary economic unit is a person or a couple who live together and all the other people who live in the same household who are financially dependent on them.

[1] Theoretically, we just require the instantaneous utility function with $u' > 0$, $u'' < 0$, $u''' > 0$. In the case of CRRA, one-period function can be defined as $u(c_t, e_{t+1}) = \frac{c_t^{1-\eta}-1}{1-\eta} + \chi \frac{[(1-\tau)e_{t+1}]^{1-\eta}-1}{1-\eta}$ where $\chi > 0$ denotes the investment motive intensity.

[2] It can also be explained as competitive firms have the same production function $F(h_t) = h_t$. In reality, we only need the function strictly increasing human capital and is homogeneous of degree one.

Equation 3 is the household's period budget constraint reflecting private credit market incompleteness. The absence of any intertemporal trade is clearly an oversimplified (but quite common) presentation of asset market incompleteness; it represents the main price of analytical tractability in the model. [3] With the absence of a credit market to finance human capital investment, requiring both consumption $c(h)$ and private education expenditure $e(h)$ to come out of disposable income.

Equation 4 is human capital production technology, describing a child's human capital. h is the product of three inputs: innate ability, stable quality of the home environment and educational investment. Innate ability reflects a random inter-generation human capital inheritance rate θ from parents. The only source of additive idiosyncratic shock θ is i.i.d. across households and time and is drawn from normal distribution $\theta \sim N\left(0, \sigma^2\right)$. As for education, all materials like teacher time, classrooms, books or computers are counted into two categlories, 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. We impose a functional form on the human capital production through education process with a constant endowment value w. We assume human capital's social discount rate is δ. The scalar σ represents the undiversified idiosyncratic investment risk due to market incompleteness (with $\sigma = 0$ corresponding to the complete market).

$$F_e\left(e(h), T\right) = w + \gamma_1 e(h)^{\epsilon_1} + \gamma_2 T^{\epsilon_2}.$$

The budget constraint can be rewritten in a way that shows how the household's optimization problem is basically a standard intertemporal portfolio choice problem where the saving policy function $s(.)$ represents the net human capital generated through all kinds of education investment minus depreciation, i.e. the optimally chosen drift of human capital [1]

$$dh = s(h)dt + h\sigma dZ \tag{5}$$

$$\begin{aligned} s(h) &= F_e(e(h), T) - \delta h \\ &= w + \gamma_1 e(h)^{\epsilon_1} + \gamma_2 T^{\epsilon_2} - \delta h \\ &= w + \gamma_1\left((1-\tau)h - c(h)\right)^{\epsilon_1} + \gamma_2 T^{\epsilon_2} - \delta h \end{aligned} \tag{6}$$

[3] Private credit market incompleteness can be simply interpreted by the fact that children cannot be held responsible for the debts incurred by their parents, as in Bénabou (2002).

[1] If deriving from the discrete version, we can take the limit $\Delta t \to 0$ and get wealth motion in the continuous-time analog of the Kesten drift-diffusion process. When we only consider the stationary case, the time index t is omitted when no confusion is incurred. A drift-diffusion process is an economic measure X_t is assumed to evolve according to $dX_t = \phi(X_t)dt + \psi(X_t)dZ_t$ where $\phi(X_t)$ is the general drift process and $\psi(X_t)$ is the general diffusion term while dZ_t represents the increment to a Wiener process. Chan et al. (1992) specify the general drift with a mean reversion speed and a long-run mean. In our model, the general drift process is embodied as $\phi(h_t) = s(h_t)$, $\psi(h_t) = \sigma h_t$

The social aggregate output reads

$$Y = \int_{\underline{h}}^{\infty} y(h)f(h)dh. \tag{7}$$

The general government balanced-budget constraint reads

$$T(\tau) = \tau Y \tag{8}$$

In the state-independent tax-transfer model, the total tax revenue from income taxation is also the lump-sum transfers. The constant public education transfer for every household exactly is equal to tax revenue, thus the equilibrium total public education expenditure solves

$$T(\tau) = \tau H(\tau) = \tau \int_{\underline{h}}^{\infty} hf(\tau, h)dh \tag{9}$$

2.2 Tax-Education Subsidy Model

In reality, the distribution of human capital is very unbalanced, which is not conducive to social progress and development. Government usually prefers subsidized public education schemes rather than state-independent transfer payments. Here we assume that the government adopts a non-linear 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. That is, if a household has a low level of income/human capital, the corresponding public education investment will be more. These non-linear subsidized redistribution policies can be interpreted as "poverty alleviation through education". 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.

In our model, there are two parts in public education. Government still offers a constant payment $\pi < \tau \int_{\underline{h}}^{\infty} hf(h)dh$ while there is also a subsidised transfer payment decreasing with household's income-level.[1]The constant part in total public education expenditure represents the degree of equalization of school resources, which can be thought of as the standard compulsory education for all citizens. The regressive subsidy part in total public education expenditure represents the average rate of education subsidization, reflects a policy adjustment made by the gadministration to ensure the fairness of education.

[1]As before, utility-maximization problem 1, social output 7 and income taxation τ remain unchanged.

Particularly, we assume 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 $\tilde{T}(h,\tau)$ for individuals reads

$$\tilde{T}(h,\tau) = \pi + \beta(\zeta)h^{-\zeta}, \quad b > 0 \tag{10}$$

The government budget constraint reads

$$\int_{\underline{h}}^{\infty} \left(\pi + \beta(\zeta)h^{-\zeta}\right) f(h)dh = \tau \int_{\underline{h}}^{\infty} hf(h)dh \tag{11}$$

$$\beta(\zeta) = \tau \frac{\int_{\underline{h}}^{\infty} hf(h)dh - \pi}{\int_{\underline{h}}^{\infty} h^{-\zeta} f(h)dh} \tag{12}$$

The saving function reads[1]

$$s(h) = \gamma_1 \left((1-\tau)h - c(h)\right)^{\epsilon_1} + \gamma_2 \left(\pi + \beta(\zeta)h^{-\zeta}\right)^{\epsilon_2} - \delta h + w \tag{13}$$

3 The Mean-Field-Games (MFG) System

In this section, HJB equation 14 and KF equation 18 characterize our MFG system.

MFG models describing the limiting behavior of stochastic differential problems, as the number of players tends to infinite, have been recently introduced by J-M. Lasry and P-L. Lions. As HACT(2020) shows, when recast in continuous time, heterogeneous agent models can be boiled down to a MFG system of two coupled partial differential equations. The first is a Hamilton-Jacobi-Bellman (HJB) equation for the optimal choices of a single atomistic individual who takes the evolution of the distribution. The second is a Kolmogorov Forward (KF) equation characterizing the evolution of the distribution, given optimal choices of individuals. Here we take the more complex tax-education subsidy model as the demonstrative case. the MFG system can be used to solve the education subsidy in the next section analogously. Keeping consumption and saving policy function defined as before, define household's value function in continuous time as $v(h) = \max\limits_{c(h)} E_{t_0} \int_{t_0}^{\infty} e^{-\rho(t-t_0)} \frac{c(h)^{1-\eta}}{1-\eta} dt$, where t_0 denotes the starting time. The domain of the following differential equations 14 and 18 is $(\underline{h}, \infty)$ where $\underline{h}$ is the borrowing limit.

The HJB equation reads

[1] $0 < \epsilon_1, \epsilon_2 \leq 1$ still capture the education elasticities.

$$\rho v(h) = \max_{c(h)} u(c(h)) + v'(h)s(h) + \frac{\sigma^2 h^2}{2} v''(h) \tag{14}$$

written with the Hamiltonian function $H(.)$, [2]

$$\rho v(h) = \max_{c(h)} H(v'(h)) + v'(h)\left[\gamma_2 w + (\pi + \beta(\zeta)h^{-\zeta})^{\epsilon_2} - \delta h\right] + \frac{\sigma^2 h^2}{2} v''(h). \tag{15}$$

Let $\frac{dH(v'(h))}{dc(h)} = 0$. The first-order condition (F.O.C.) solves [1]

$$\left[(1-\tau)h - e(h)\right]^{-\eta} = v'(h)\gamma_1 \epsilon_1 e(h)^{\epsilon_1 - 1} \tag{16}$$

or

$$c(h)^{-\eta} = v'(h)\gamma_1 \epsilon_1 \left[(1-\tau)h - c(h)\right]^{\epsilon_1 - 1} \tag{17}$$

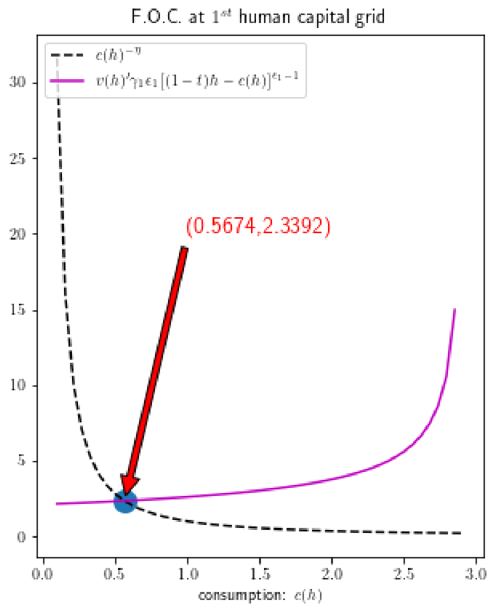


Figure 1: The Poorest

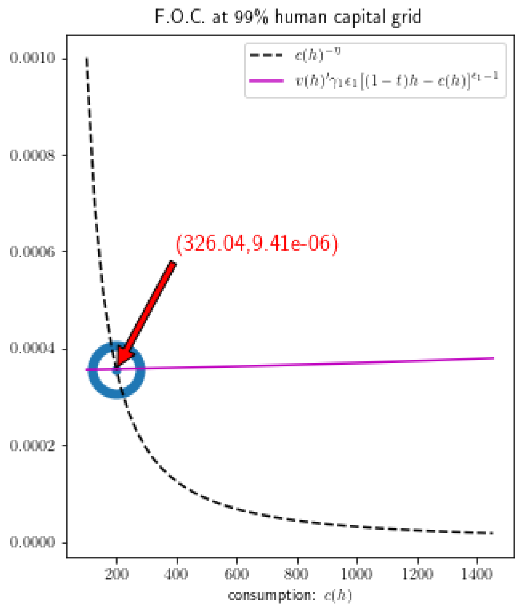


Figure 2: The Richest

Although the F.O.C. is non-linear w.r.t. $c(h)$, there exists a unique solution. [2] As we all know, people desire to translate their consumption from periods of high income to

[2]Hamiltonian function $H(v'(h)) = \max_{c(h)}\{u(c(h)) - v'(h)\gamma_1\left[(1-\tau)h - c(h)\right]^{\epsilon_1}\}$.

[1]F.O.C. 16 & 17 can also be solved by $0 = -\frac{du(.)}{de} + v'(.)\frac{ds(.)}{de}$ through HJB equation 14. It can also be noted as $u'(c) = v'(h)\gamma_1\epsilon_1\left[(1-\tau)h - c(h)\right]^{\epsilon_1 - 1}$ which reflects the relationship between $u'(c, h, v'(h))$ and $v'(h)$.

[2]The algorithm speed may slow down because we're using a nonlinear processor, this problem can be mitigated by using parallel computing.

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. [1] An example for policy functions for consumption, private education investment and saving are illustrated in figure 3 and 4.

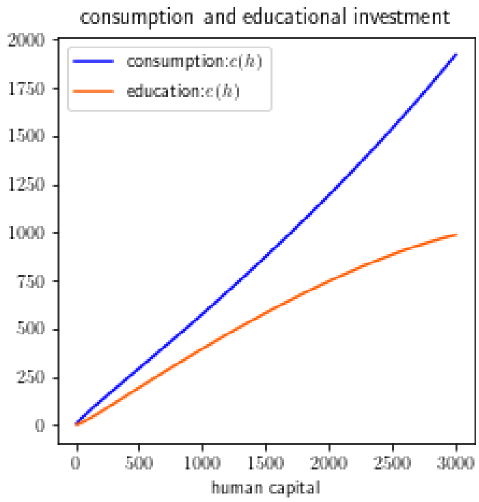


Figure 3: Consumption&Education

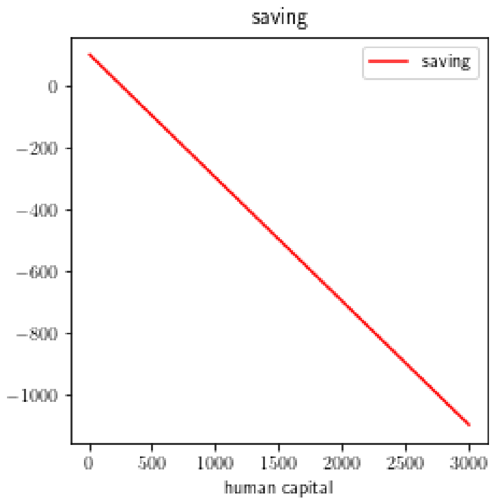


Figure 4: Saving

Let $f(.)$ denotes the density function of stationary distribution which not change over time. The KF equation satisfies [2]

$$0 = -\frac{\partial}{\partial h}\big(s(h)f(h)\big) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial h^2}\big(h^2 f(h)\big) \tag{18}$$

written in the first-order linear differential equation form as

$$\frac{df(h)}{dh} + p(h)f(h) = q(h) \tag{19}$$

$$\begin{cases} p(h) = 2\big[\frac{1}{h} - \frac{s(h)}{\sigma^2 h^2}\big] \\ q(h) = \frac{2C_0}{\sigma^2 h^2} \end{cases} \tag{20}$$

[1]In this particular setting, highest human capital value is 3,000. Consumption of the household with highest income (326.04) is not too large when compared with the median of human captial, the F.O.C. consumption solves 94.9323, which can be seen in APPENDIX figure 23.

[2]Its time-dependent counterpart reads $\frac{\partial f(h_t)}{\partial t} = -\frac{\partial}{\partial h_t}\big(s(h_t)f(h_t)\big) + \frac{\sigma^2}{2}\frac{\partial^2}{\partial h_t^2}\big(h_t^2 f(h_t)\big)$.

with boundary conditions in proposition 3.1.

Proposition 3.1 *Kolmogorov Forward equation 18 satisfing*

$$1 = \int_{\underline{h}}^{\infty} f(h)dh \tag{21}$$

$$0 = \int_{\underline{h}}^{\infty} s(h)f(h)dh \tag{22}$$

Condition 21 is the natural property of density function. [1] Condition 22 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 19 is advantageous to use iterative algorithms, we develop an iterative algorithm for solving human capital density functions in the following pages. A further explanation of the distribution patterns and proofs will be given in Section 4.3.

4 Propositions of Household's Behavior and Human Capital/Income Distribution

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-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

HJB equation is defined on $(\underline{h}, \infty)$, but in practice it sometimes helps numerical stability to impose a state constraint

[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.

$$\underline{h} \leq h \leq h_{\max} \tag{23}$$

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, $s(\underline{h}) > 0$.*

To support corollary 4.1, We need a "total employment hypothesis", which means all households in our model are in employment, referring to the high-earning state in HACT(2020).

Saving at lower boundary $\underline{h}$ is strictly positive. Then a state constraint boundary condition needed is $s(\underline{h}) > 0$. Since F.O.C of HJB equation $u'(c(\underline{h})) = v'(\underline{h})\gamma_1\epsilon_1[(1-\tau)\underline{h} - c(\underline{h})]^{\epsilon_1 - 1}$ still hold at $h = \underline{h}$, an alternative condtion 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, hus the KF equation requires no boundary condition at $\underline{h}$. The positive saving derived from "total employment hypothesis" $s(\underline{h}) > 0$ causes the lower bound of the human capital space $h = \underline{h}$ acting as a reflecting barrier of the human capital accumulation process $\{h(t)\}_{t=0}^{\infty}$. It would not be stuck at $\underline{h}$ and has a non-degenerating stationary distribution.

In short, the saving and consumption behaviours of the low-income (at lower boundary of human capital/income) reflect the precautionary motive. As set in our model, the household suffers from idiosyncratic investment risk generated from the Brownian motion, which causes the precautionary saving motive.

4.2 Behaviors of the High-Income Households

We will give some propositions about the high-income household's consumption and saving behavior when human capital reaches to an infinitely large value.

Lemma 4.1 *Take the derivative of HJB equation 14 on h of both sides (We use the Envelope theroem) to get a general form of Euler equation*

$$(\rho - s'(h))v'(h) = u'(c)c'(h) + v''(h)\big(s(h) + \sigma^2 h\big) + v'''(h)\frac{\sigma^2 h^2}{2} \tag{24}$$

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.

Euler equation reads

$$\rho = \gamma_1\big(1-\tau\big) - \delta + (a-1)\vartheta + a(a-1)\frac{\sigma^2}{2} < 0 \tag{25}$$

Consumption and saving satisfying

$$\frac{s(h)}{h} \sim \vartheta = -\frac{\rho+\delta}{\eta} + \frac{\gamma_1\big(1-\tau\big)}{\eta} - (1-\eta)\frac{\sigma^2}{2} \tag{26}$$

$$\frac{c(h)}{h} \sim \phi = \frac{(1-\tau)(\eta-1)}{\eta} + \frac{(1-\eta)\frac{\sigma^2}{2} - \delta}{\gamma_1} + \frac{\rho+\delta}{\gamma_1\eta} \tag{27}$$

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

$$\rho = -\delta + (a-1)\vartheta + a(a-1)\frac{\sigma^2}{2} \tag{28}$$

Consumption and saving satisfying

$$\frac{s(h)}{h} \sim \vartheta = -\frac{\rho+\delta}{1-a} - \frac{\sigma^2 a}{2} = \frac{\rho+\delta}{1-\eta-\epsilon_1} - \frac{2-\eta-\epsilon_1}{2}\sigma^2 \tag{29}$$

$$\frac{c(h)}{h} \sim \phi, \quad \phi^{-\eta} = (2-\eta-\epsilon_1)B\gamma_1\epsilon_1[1-\tau-\phi]^{\epsilon_1-1} \tag{30}$$

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

$$\begin{aligned}\frac{s(h)}{h} &\sim \vartheta = -\delta \\ \frac{c(h)}{h} &\sim \phi = 1-\tau\end{aligned} \tag{31}$$

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.*

To sum up, as $h- > \infty$, negative saving function $s(h) < 0$ is asymptotically linear on h, which induce the mean-reverting process of human capital. When $\vartheta < 0$, the consumption function $c(h)$ is asymptotically linear on h and positive. Moreover, because only monotone increasing and concave properties of CRRA function were used in the proof process, we can extend the properties to other utility functions. [1] 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

One of the HACT's main theoretical result is that an analytic solution to the KF equation characterizing the stationary distribution for given saving functions. We can show that if individuals have the constant relative risk aversion (CRRA) utility, the stationary human capital distribution has a Pareto tail when $\zeta \leq 1$, and derive an analytic expression for the tail parameter. Given the numerical solution of $c(h)$ from HJB numerical solutions, human capital distribution $f(h)$ can be solved throught the KF equation. If absolute risk aversion remains finite, $\underline{R} = -\frac{u''(c(\tilde{\underline{h}}))}{u'(c(\tilde{\underline{h}}))} < \infty$ and relative risk aversion is bounded above for all c, $\eta = -\frac{cu''(c)}{u'(c)} < \infty$, then there exists a stationary equilibrium (HACT(2020)) . The human capital distribution is smooth or the density of human capital is continuous and differentiable for all h.

Reconsider proposition 4.1, we should notice that the condition $s(\underline{h}) > 0$ must hold. If $s(\underline{h}) = 0$ it will turn to the unemployed type people remarked in HACT(2020) whose distribution has a Dirac mass point.[2] In our model, the stationary distribution does not have a Dirac point mass, i.e. its CDF satisfies $F(\underline{h}) = 0$ or just $g(h) < \infty$ for all h. Thus $s(\underline{h})f(\underline{h})$ equals some negative constant and $s(h)f(h)$ equals to some constant for all h.

[1]Benhabib, Bisin and Zhu(2015) has shown that in continuous time.

[2]HACT proves that the stationary distribution of the low-income type has a Dirac point mass at the borrowing constraint, i.e. its CDF satisfies $F(\underline{h}) = m > 0$ or as $h \to \underline{h}$, $f(h)$ explodes.

And from proposition 4.1, we can also claim that the support of the stationary human capital distribution is bounded above at some $h_{\max} < \infty$. To sum up, it does not have a mass point at both $\underline{h}$ and $h_{\max}$.

To sum up, There are two corollaries leading to the the stationary human capital distribution.

Corollary 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 $h_{\max}$.*

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

$$\bar{F}(h) = \Pr(H > h) = \begin{cases} \left(\frac{h_m}{h}\right)^{\alpha} & h \geq h_{\mathrm{m}} \\ 1 & h < h_{\mathrm{m}} \end{cases} \tag{32}$$

In this definition, h_m (scale parameter) is the minimum possible value of H and α is a positive constant (shape parameter). Furthermore, in the context of distribution, α is named the Pareto Index (Tail Index) here 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 wealth/income exceeding a certain level in the sense that it must have a tail of order $(\frac{1}{h})^{\alpha}$.

Assumption 4.2 *We assume that*

when $\epsilon_1 = 1$:

$$2(\rho + \delta) > 2\gamma_2(1 - \tau) + \sigma^2\eta(\eta - 1), \tag{33}$$

when $2 - \eta < \epsilon_1 < 1$:

$$2(\rho + \delta) > \sigma^2(2 - \eta - \epsilon_1)(1 - \eta - \epsilon_1). \tag{34}$$

Assumption 4.2 is to guarantee the negativity of saving function as human capital goes to infinite, i.e. $\bar{s} = \lim_{h\to\infty} \frac{s(h)}{h} < 0$, which is a necessary condition of having the income stationary distribution.

Recall our endogenous human capital accumulation process

$$dh(t) = (h)dt + \sigma h dB(t).$$

Private firms suffer from a Browian idiosyncratic investment risks. Idiosyncratic investment risks induce the precuationary savings, and it would push the $\bar{s}$ to be positive. The higher η and σ are, the stronger the households' precuation savings motive is. Thus, the upper bound $2(\rho+\delta)$ confines the households' savings motive. And again, it underlines the necessary of a high-income assumption 4.1.

More generally, we can make a conclusion. We claim that if households have CRRA utility, the stationary distribution has a Pareto tail with an analytic expression as in proposition 4.2.

Proposition 4.2 *Under CRRA utility assumption, "negative saving assumption" 4.2 and "high wage assumption" 4.1, if* $2-\eta < \epsilon_1 \leq 1$, *there is a unique stationary wealth distribution which follows an asymptotic power law, that is the p.d.f. of human capital* $f(h)$ *exibits a Pareto tail*

$$f(h) \sim \epsilon h^{-\Theta-1}, \quad h \to \infty \tag{35}$$

where

$$\begin{aligned} \epsilon &= \frac{2}{\sigma^2} \lim_{h\to\infty} \ln\left(\frac{y(h)}{h^{\mu}}\right) \\ \Theta &= 1 - \frac{2\bar{s}}{\sigma^2} \\ \mu &= \frac{2\bar{s}}{\sigma^2}, \quad \bar{s} = \bar{s}_a, \bar{s}_b \end{aligned}$$

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

Let $x = \ln h$, *we have proposition*

$$\frac{\ln f(x)}{x} \sim \epsilon\Theta, \quad h \to \infty \tag{36}$$

This result extends Benhabib, Bisin, and Zhu (2015)[9] and Proposition 10 in HACT(2020) from a partial to a general equilibrium setting. Top wealth inequality $\frac{1}{\Theta}$ is increasing in volatility σ, risk aversion η, and decreasing in the rate of time preference ρ. It is oblivous that $\bar{s}_a > \bar{s}_b$ and furthermore, $\frac{1}{\Theta_a} > \frac{1}{\Theta_b}$.

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 ($\epsilon_1 = 1$) is fatter than that under non-linear private education return ($\epsilon_1 < 1$). Moreover, linear taxation and the introduce 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. And in our following calibration, we find that the income taxation would influence the gini coeffcient of income distribution.

A fact about Pareto tail is that if a variable h follows a Pareto distribution $f(h) \propto h^{-\Theta-1}$, then its logrithm $x = \log h$ follows an exponential distribution $f(x) \propto e^{-\Theta x}$ and hence $\log f(x)$ is a linear function of x where the slope equals the tail exponent Θ. We use this property in the text to examine the Pareto tail. We are going to show check out the distribution behaves asymptotically like a Pareto distribution by showing that the logarithm of the density of log human capital $\ln f(x)$ is asymptotically linear in the logarithm of human capital $x = \ln h$.

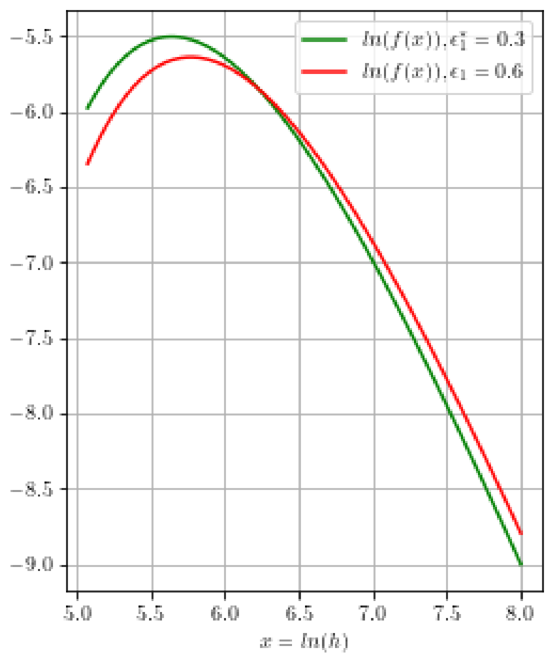


Figure 5: Check the Pareto Tail

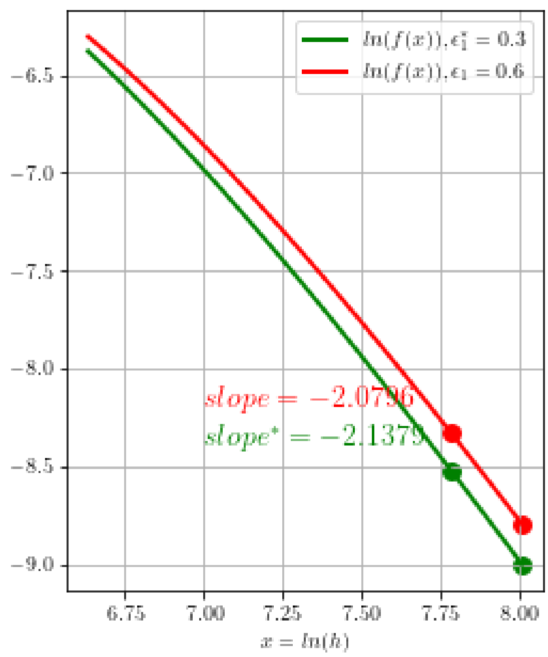


Figure 6: Compare Tails with Different ϵ_1

When $\tau^* = 0.0128\%$, the tail exponent of stationary distribution looks like figure 5, whose slope is around -2. The tail exponent of stationary distribution as the private education elasticity changes is shown as in figure 6. As we can see from the above graph, an incremental decrease in the elasticity of private education investment ϵ_1 leads to an incremental increase in the Pareto tail exponent of the stationary distribution of income capital, 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.

4.3.2 Inverse Gamma Pattern

In particular, we were able to prove that the distribution obeies a specific Inverse Gamma pattern under which we set it up. The general solution of KF equation 18 can be written as a variable upper bound integral form as we described in the following propsition 4.3.

Proposition 4.3 *Based on KF equation 18 and its first-order linear differential form 19, the p.d.f. of human capital can be written as (first-order differential equation of variable upper bound integral form)*

$$f(h) = C_1 e^{-2\int_{\underline{h}}^{h} p(x)dx} + C_2 \int_{\underline{h}}^{h} x^{-2} e^{-2\int_{x}^{h} p(z)dz} dx \tag{37}$$

$$\begin{cases} p(h) = 2[\frac{1}{h} - \frac{s(h)}{\sigma^2 h^2}] \\ q(h) = \frac{2C_0}{\sigma^2 h^2} \end{cases} \tag{38}$$

Notice that C_0, C_1 and C_2 are constants. From the boundary conditions in proposition 3.1, we can prove that $C_2 = 0$ and $f(h)$ follows an Inverse Gamma distribution as

$$f(h) = C_1 e^{-\frac{\beta}{h}} h^{-\alpha-1} \tag{39}$$

$$C_1 = \frac{1}{\int_0^{\infty} e^{-2\int_{h_0}^{h} \frac{1}{x} - \frac{s(x)}{\sigma^2 x^2} dx} dh} \tag{40}$$

For simplicity, let's give some notes

$$
\begin{cases}
g(h) = e^{-\int_{h_0}^{h} p(x)dx} = e^{-\int_{h_0}^{h} \frac{2}{x} - \frac{2s(x)}{\sigma^2 x^2} dx} = (h^{-2} - h_0^{-2}) e^{\int_{h_0}^{h} \frac{2s(x)}{\sigma^2 x^2} dx} \\
z(x) = x^{-2} e^{\int_{h_0}^{x} p(z)dz} = x^{-2} e^{\int_{h_0}^{x} \frac{2}{z} - \frac{2s(z)}{\sigma^2 z^2} dz} \\
G = \int_0^\infty g(h)dh \\
G_h = \int_0^\infty hg(h)dh \\
Z = \int_0^\infty g(h) \int_1^h z(x)dxdh \\
Z_h = \int_0^\infty hg(h) \int_1^h z(x)dxdh
\end{cases} \tag{41}
$$

Let's substitute the expression in proposition 4.3 of human capital density into boundary conditions and get the condition system as in equation 77

$$
\int_0^\infty f(h)dh = 1
$$
$$
\int_0^\infty s(h)f(h)dh = 0
$$

or an equation system

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

According to the KF equation 18, $C_2 = \frac{2C_0}{\sigma^2}$ and $C_0 = f(\underline{h})$. Although from the numerial method, C_2 is calculated to be zero, we still want to prove it from a mathematical methods. The proof conditions include constant ϑ in proposition 4.1 less than zero. We can prove G and G_h solves certain constants as $h- > \infty$ and Z_h and Z goes to infinite. Based on this, we claim that

$$
C_2 \sim 0, \quad h \to \infty. \tag{42}
$$

See detailed **Proof** of Inverse Gamma distribution in *APPENDIX D* . Figures 7 and 8 shows us how the human capital distribution changes with tax rate or private education investment return increasing.

We find in these figures that different shocks influence the different parts of the human capital distribution. 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

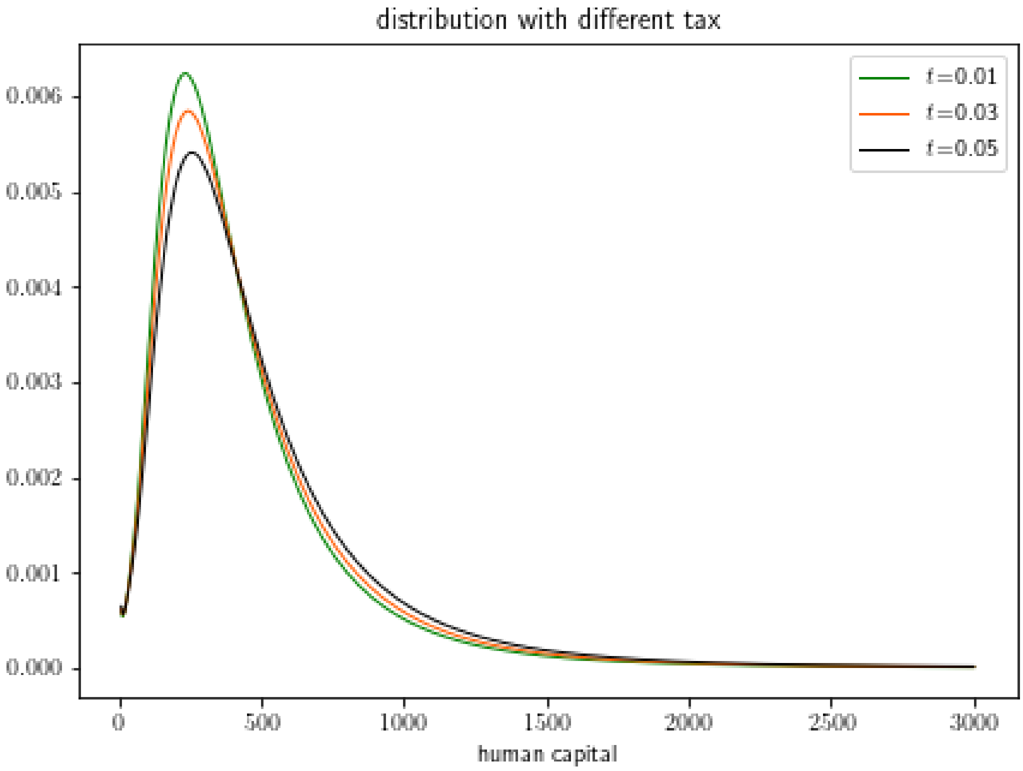


Figure 7: Change with τ

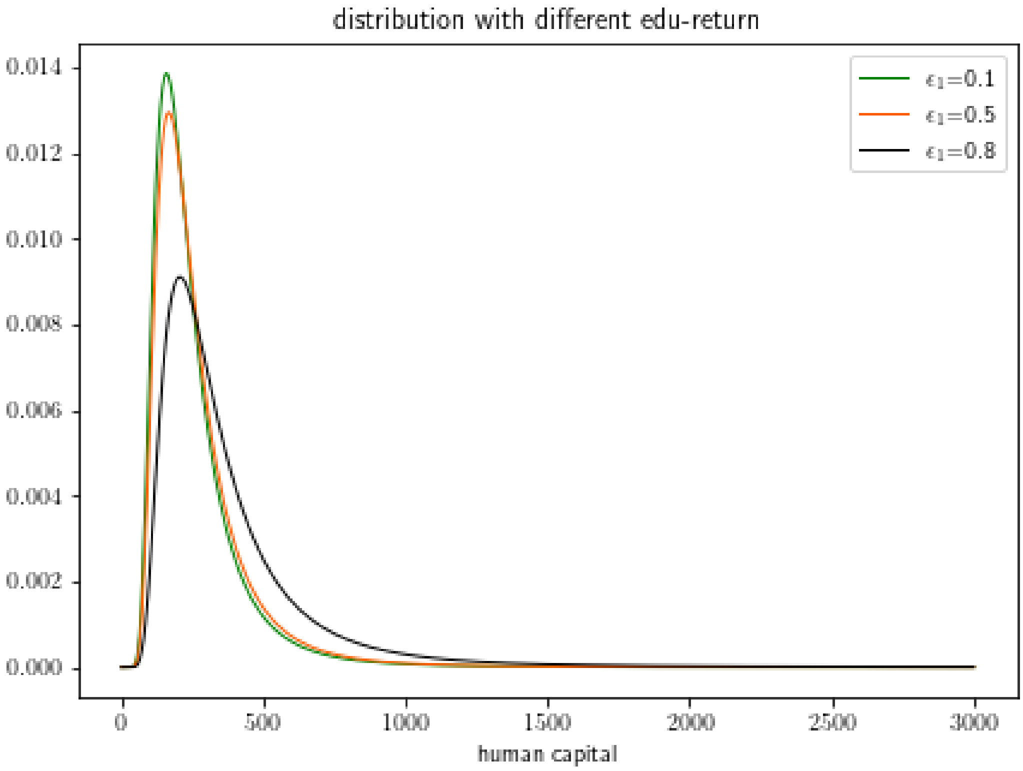


Figure 8: Change with ϵ_1

investment risk, as emphasized by Park (2020)[35] which studies optimal tax scheme that the scheme depends on the sources of income inequality.[1]

5 Computation of Stationary Human Capital Distribution

In this section, we'll show you some mathematical results and introduce algorithms solving MFG system and equilibrium. During computation, we develop a iterative method to generate human capital distribution apart from original algorithm in HACT(2020) and a bisection method is used to get the balanced government transfer payments.

5.1 Finite Difference (FD) Algorithm

Numerical methods for the approximation of the stationary and evolutive versions of MFG models have been proposed by the Achdou e.t. al in previous works. Convergence theorems for these methods are also proved under various assumptions. Among these methods, we chose the finite difference (FD) method based on work by Achdou and Capuzzo-Dolcetta (2012)[2], Achdou(2013)[3], and HACT(2020). [2] Apart from their work, we add a balanced government budget constraint to our criteria for determining whether the distribution of wealth has reached stationary distribution. Specially, for approximating density f, we also develop an iterative algorithm proposed in the part 4.3.

The key idea is that FD method transforms our system of differential equations into a system of sparse matrix equations. The core work is to estimate the derivative of the value function. We approximate the derivative $v'(h)$ with either a forward or backward difference approximation

$$v'\left(h_i\right) \approx \frac{v_{i+1}-v_i}{\Delta h} =: v'_{i,F} \quad \text{or} \quad v'\left(h_i\right) \approx \frac{v_i-v_{i-1}}{\Delta h} =: v'_{i,B}$$

The second derivative is

$$v''\left(h_i\right) \approx \frac{v_{i+1}-2v_i+v_{i-1}}{(\Delta h)^2}.$$

It's important whether and when a forward-backward difference approximation is used.

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

[2]See *https://nbviewer.org/github/QuantEcon/QuantEcon.notebooks/blob/master/aiyagari_continuous_time.ipynb.* We made some adjustments to the lower&upper boundary of the original intensity matrices.

The solution is to use so-called "upwind scheme." The key idea is to use a forward difference approximation whenever the drift of the state variable is positive and to use a backward difference whenever it is negative.

With this goal in mind, we approximate both v and f at I discrete points in the space dimension, $h_i, i = 1, \ldots, I$ Denote the value function and distribution along this discrete grid using the vectors $\mathbf{v} = (v(h_1), \ldots, v(h_I))^{\mathrm{T}}$ and $\mathbf{f} = (f(h_1), \ldots, f(h_I))^{\mathrm{T}}$; both $\mathbf{v}$ and $\mathbf{f}$ are of dimension $1 \times I$, the total number of grid points in the individual state space. The end product of our discretization method will be the following system of matrix equations:

$$\rho \mathbf{v}^{n+1} = \mathbf{u}^n(\mathbf{v}) + \mathbf{A_H}^n(\mathbf{v})\mathbf{v}^{n+1} \tag{43}$$

$$\mathbf{0} = \mathbf{A_K}(\mathbf{v})^{\mathrm{T}}\mathbf{f} \tag{44}$$

The first equation is the discretized HJB equation 14, the second equation is the discretized KF equation 18. The $I \times I$ matrix $\mathbf{A_H}(\mathbf{v})$ has the interpretation of a transition matrix that captures the evolution of the idiosyncratic state variables in the discretized state space. It turns out to be extremely sparse. $\mathbf{A_K}(\mathbf{v})$ in the second equation denotes the other matrix designed for KF equation.

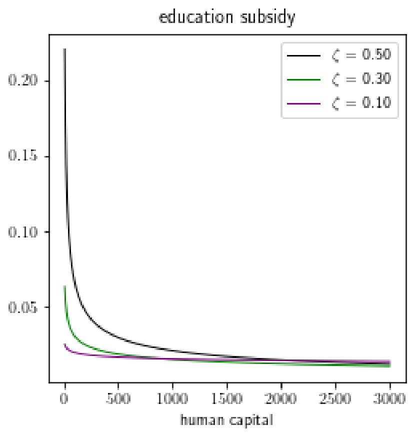


Figure 9: Subsidy

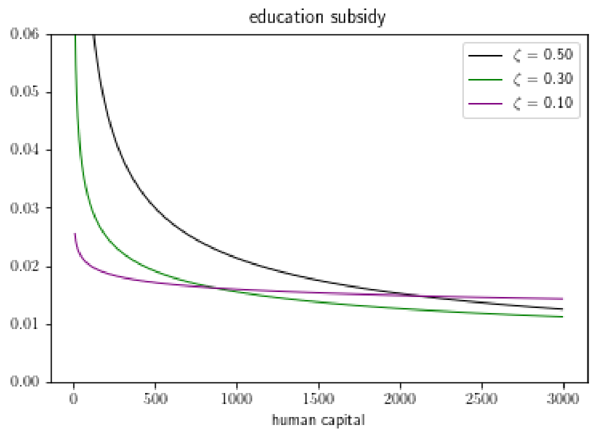


Figure 10: Changes with Different ζ

There are some comparative static analysis for policy functions and value function from numerical solutions in figures 9, 10, 11, 12, 14 and 13.

We developed two kinds of numerical method to generate the human capital distribution. One is called "FD method" solving KF equation solution in MFG-system, the other named "iterative method" KF equation solution. Among them, FD method is derived from the original numerical solution designed for Aiyagari model (1994) with labor-income risk (Two-stage Poisson process) and investment risk (Brownian process) in continuous time with some revamps on intensity matrix $\mathbf{A_K}$. In the process of applying

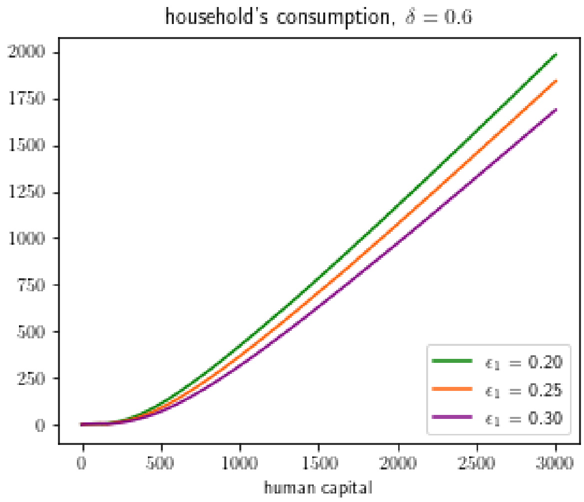


Figure 11: Consumption with Different Education Elasticity

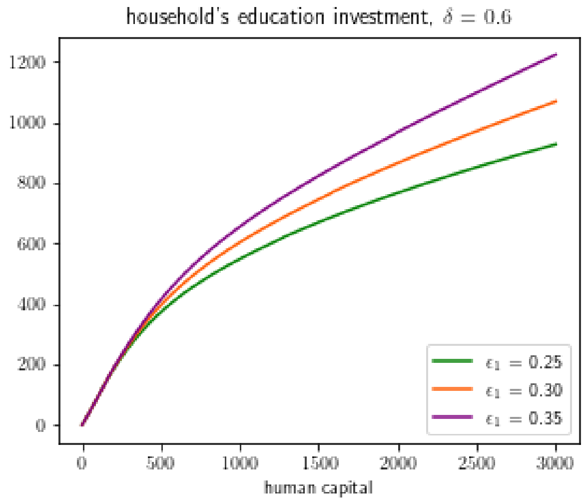


Figure 12: Education Investment with Different Education Elasticity

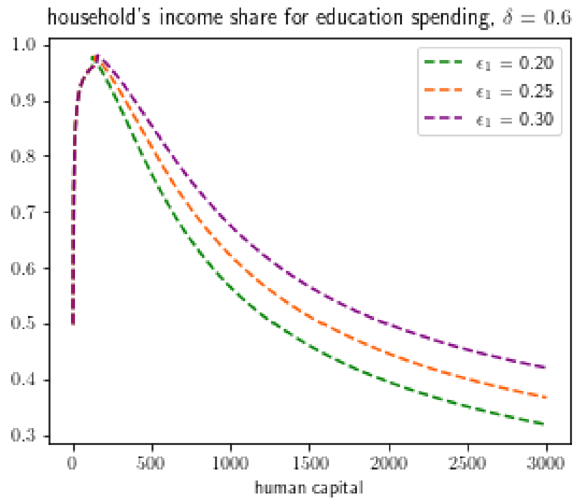


Figure 13: Education Sharing with Different Education Elasticity

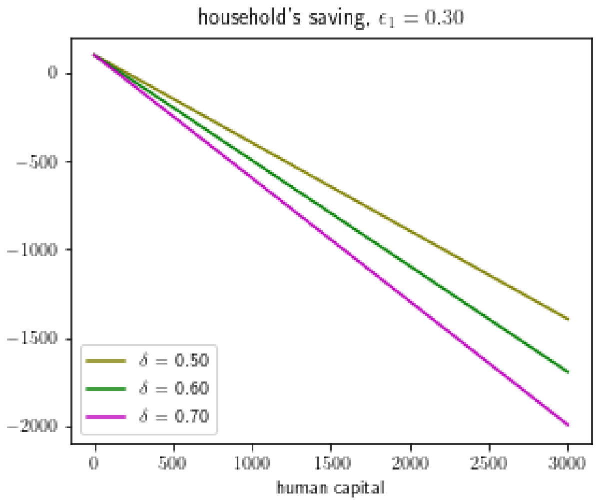


Figure 14: Saving with Different Depreciation

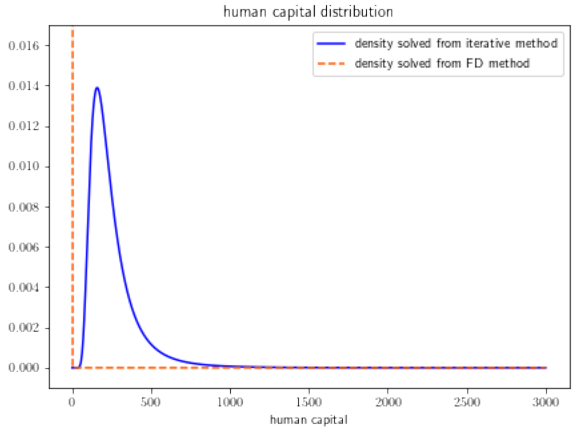


Figure 15: with mass point

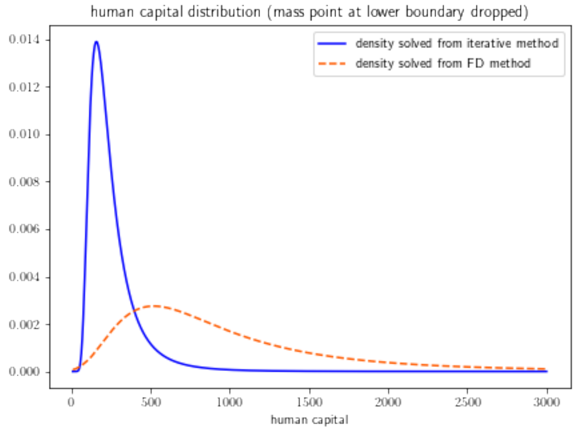


Figure 16: mass point dropped

this 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. We can simply drop out this outliers, but this treatment will result that the tail of the distribution is not very fat. Thus, we develop an "iterative" KF equation solution. Because our KF equation has no Poisson process, we can easily write it into the ODE form described above, then apply to discretization and iteration. The density function drawn in this way has no mass point and has a fatter tail. A comparison of the results obtained by the two methods is shown in figure 16.

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 and try to replicate the income inequality in the U.S. economy.

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	$\zeta = 0.50$

In our calibration model, we divide the parameters into two groups as in table1. 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.[1] In their definition, income includes labor earnings and capital incomes. The volatility of Brownian motion would influnce the Top 1% income shares. The higher 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.

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, education return coefficients.

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	$s(h) = \gamma_1 e(h)^{\epsilon_1} + \gamma_2\left(\pi + \beta(\zeta)h^{-\zeta}\right)^{\epsilon_2} - \delta h + w$	$\epsilon_1 = 0.90$
		$\epsilon_2 = 0.50$
		$\pi = 0.1$
		$\zeta = 0.50$
	$c(h) + e(h) = (1-\tau)h$	$\tau = 0.128$

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. [2] 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

[1] 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).

[2] 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.

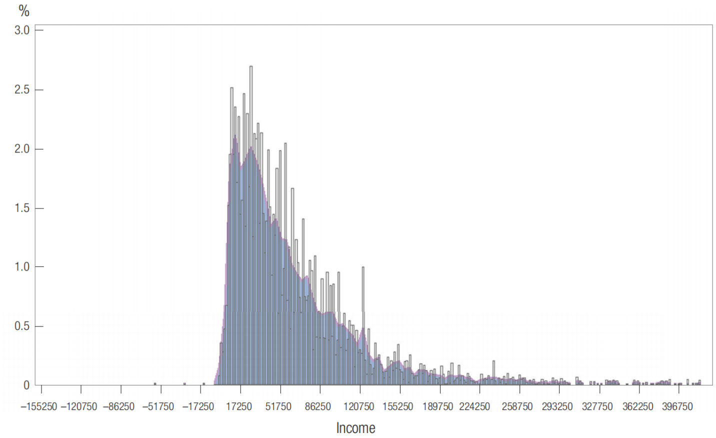


Figure 17: Income Distribution (2007 SCF, USD)

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 17 displays the histogram of the income distribution.

The income distribution of our calibrated model matches the target moments reasonably well. Table 3 displays the wealth share moments we use and compares the moments in the data with those obtained simulating the model in terms of income distribution ranking. Quantiles are values that separate fractions of the population. See the method of estimating income quantiles in *APPENDIX F*.

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).

Compared with the traditional tax-tranfer economy, the education subsidy scheme can generate a more equal human capital/income distribution. Generally, our model is more equal than 2007 SCF 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 for the human capital distribution to see their incidence on equity. Figure 18 shows how the income Gini coefficient changes with linear tax rate. Table 4 shows how the income Gini coefficient changes with education elasticity and public education scheme.

Figure 18: Gini Changes with Linear Income Taxation ($\epsilon_1 = 0.9, \sigma = 0.56$)

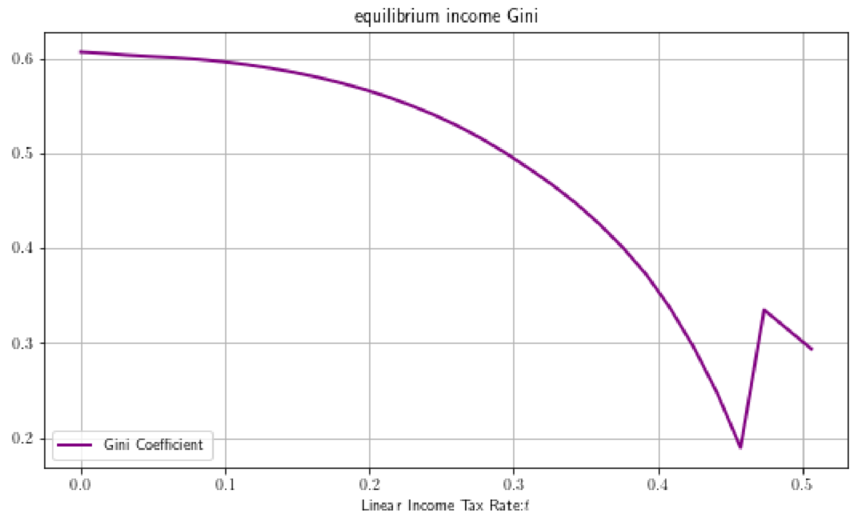
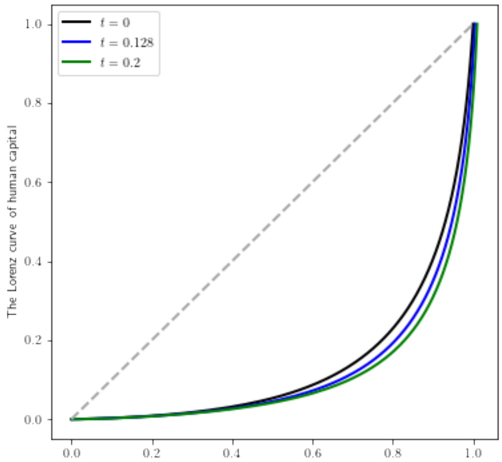


Table 4: Gini Changes of Subsidy Model

	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

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 as in figure 19. However in 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

Figure 19: Income Lorenz Curve

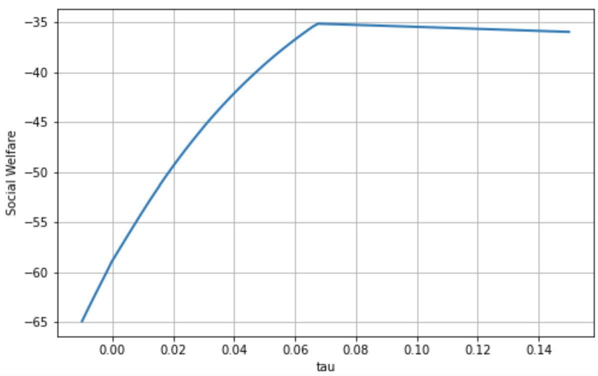


negative impact on their children's private education return. From this perspective, we also explain why it is more pronounced that Gini coefficient increases w.r.t. privated education elasticity ϵ_1 than w.r.t. public education elasticity ϵ_2. As $\epsilon_1 \to 1$, the effect will be close to linear 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.

We don't just evaluate economic policies from the perspective of income equality. Moreover, policies are most often evaluated according to some social welfare criterion like utility aggregates. Because of the concavity of individual utility, any such utilitarian criterion rises with all current and future redistributions. Define social welfare function

$$\max_{\tau,\zeta} W(\tau) = \max \int_{\underline{h}}^{\infty} v(h) f(h,\tau) dh \tag{45}$$

Figure 20: Social Welfare (Transfer Model)



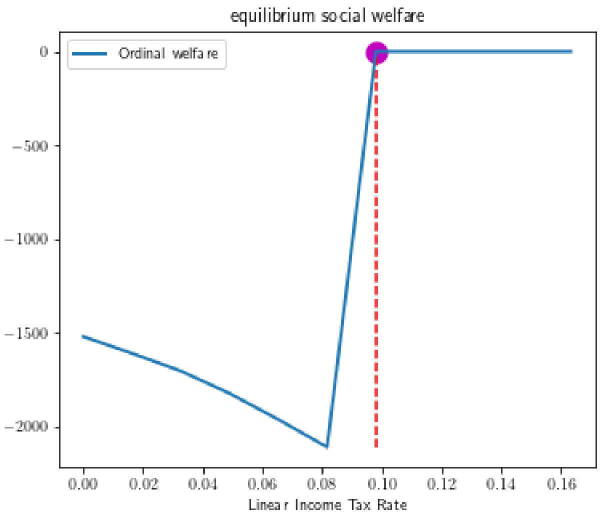


Figure 21: $\epsilon_1 = 0.9, \sigma = 0.56$

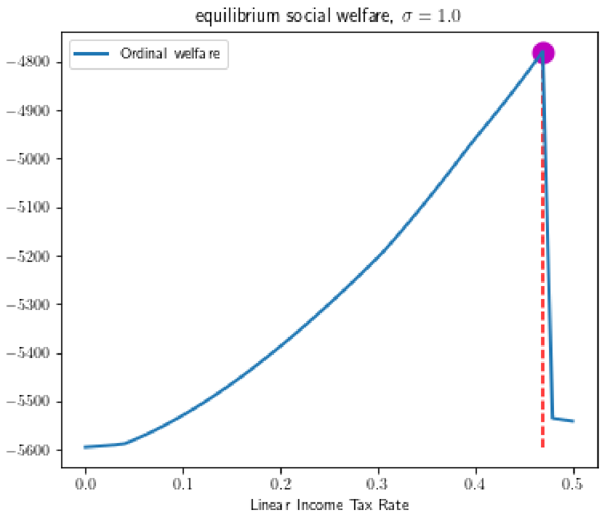


Figure 22: $\epsilon_1 = 0.9, \sigma = 1.0$

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 HJB and KF equation again, we can draw the social welfare function, and then select the optimal tax for social welfare as in figures 21,22 while for tranfer model in figure 20. And under the calibrated idiosyncratic risk $\sigma = 0.56$, the optimal tax rate is very close to the optimal rate 0.128 in Krebs(2003)[15].

Apart from social welfare, 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

[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.

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

Income distributions have been shown skewed to the right, displaying a "fat-tail" pattern with a large accumulation of mass in a relatively small range of values. As HACT(2020) shows, the capital accumulation process in a model only with idiosyncratic labor earning shocks, which has an upper bound, would raises difficulties to generate large fractions of fat tails. Incorporating an education investment risks in human capital which plays an important role in replicating the fat tail of the income distribution, our model fits well with the Top 1% group fraction in 2007 U.S. income distributions.

With the idiosyncratic investment risk represented by the volatility part of the Brownian motion, we introduce an exponential educational elasticity to our education subsidy model under linear income taxation. Compared with the traditional tax-tranfer economy, the education subsidy scheme can generate a more equal human capital/income distribution. Idiosyncratic investment risk brings more precautionary savings and more government subsidies. As for the subsidy scheme, the increase in education subsidy $\beta(\zeta)h^{-\zeta}$ and constant compulsory education π would, as expected, makes society more equal. There are both income effects and substitution effects when an increase educational elasticity occurs.

For the future work, we do social welfare decomposition. we would investigate the subsidy incidence on the economy by a perturbation method, analysing the education scheme incidence through perturbations on the MFG system.[1] Then, we want to derive the optimal subsidy formula in the tax-education subsidy model model. We will focus on the income distribution channel in incidence analysis, which is important for the pecuniary externality and the inequality effect which affects the difference across the individual marginal 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.

[1]If you are interested in it, please see future work on Github: https://github.com/yinuonino.

[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] Loury, G., 1981. Intergenerational transfers and the distribution of earnings. Econometrica 49, 843–867.

[27] Mulligan, C., 1997. Parental Priorities and Economic Inequality, University of Chicago Press, Chicago. Murphy, K., Topel, R., 1990. Efficiency wages reconsidered: theory and evidence. In: Weiss, Y., Fishelson, G. (Eds.), Advances in the Theory and Measurement of Unemployment, MacMillan, London, pp. 204–242.

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

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

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

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

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

[33] 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.

[34] 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.

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

[36] 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.

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

A Numerical Solution of the Stationary Human Capital Distribution

A.1 Solving HJB Equation: Implicit Finite Difference (FD) Method using MFG System

We use an implicit upwind method supporting arbitrarily step size Δ. Denoting the distance between the two human capital grid points by equispaced grid Δh. Notice that the HJB equations are highly non-linear, and therefore so is the system of equations has to be solved using an iterative scheme. We use a finite difference (FD) method and approximate the value function $v(.)$ at I discrete points in the human capital space dimension, $h_i, i = 1, \ldots, I$. In period 0, the household born with initial human capital $h_{i,0} > 0$ will decide how much to consume and to invest in private education $e_{i,0} \geq 0$ out of period 0 disposable income. The Household makes the allocation decision in every period t to maximize its intertemporal utility. The condition $e_t \geq 0$ is used to make saving at lower boundary positive.

Using the short-hand notation $v_i \equiv v\left(h_i\right)$ and so on.

A.1.1 Upwind Scheme

The rough idea of upwind scheme is to use a forward difference approximation whenever the drift of the state variable is positive and backwards difference whenever it is negative.

Since Brownian motion is introduced as a random term, we need the first- and second-order discrete estimators of the value function. (If Brownian motion is not included, only the first-order term is needed.)

The upwind scheme requires us to calculate the first derivative of the value function $v'(.)$ using forward and backward estimates respectively. Define $\left(v_{i,F}^n\right)'$ as the forward estimation at iteration n while $\left(v_{i,B}^n\right)'$ the backward one. Thus the first-derivative $v_i' = v'\left(h_i\right)$ respect to capital h_i is approximated with either a forward or a backward difference approximation as

$$\left(v_{i,F}^n\right)' \approx \frac{v_{i+1}^n - v_i^n}{\Delta h} \tag{46}$$

$$\left(v_{i,B}^n\right)' \approx \frac{v_i^n - v_{i-1}^n}{\Delta h} \tag{47}$$

The second-derivative $v_i'' = v_i''(h_i)$ respect to capital h_i is approximated with just the term of value function before and after it

$$(v_i^n)'' \approx \frac{v_{i+1}^n - 2v_i^n + v_{i-1}^n}{(\Delta h)^2} \tag{48}$$

Compute consumption policy according to both the backwards and forward difference approximations of value function v_i' in 47. From F.O.C. 17 (Here we let $\epsilon_1 < 1$ through the whole process), the approximation is externalized as

$$\begin{aligned} (c_{i,F}^n)^{-\eta} &= (v_{i,F}^n)'\gamma_1\epsilon_1\big[(1-\tau)h_i - (c_{i,F}^n)^{-\eta}\big]^{\epsilon_1 - 1} \\ (c_{i,B}^n)^{-\eta} &= (v_{i,B}^n)'\gamma_1\epsilon_1\big[(1-\tau)h_i - (c_{i,B}^n)^{-\eta}\big]^{\epsilon_1 - 1} \end{aligned} \tag{49}$$

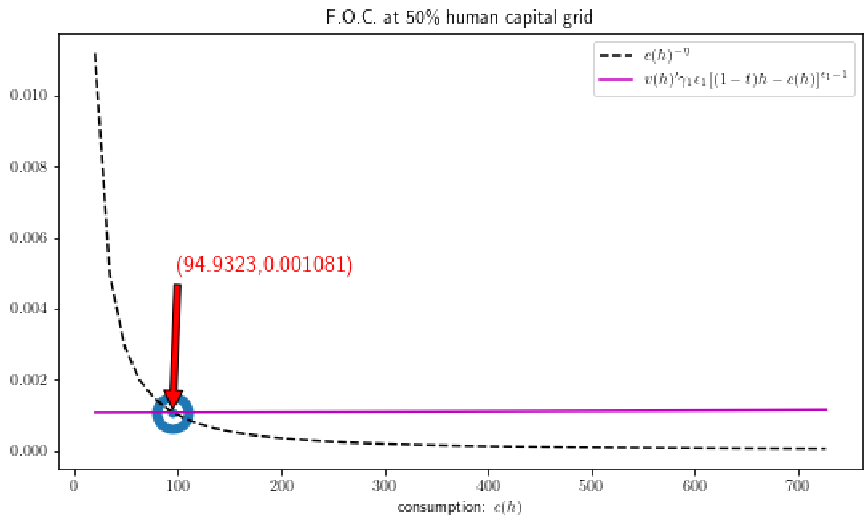


Figure 23: F.O.C solved at the median of human capital

In order to clearly check whether there is a unique solution, we draw out the fixed curve $c^{-\eta}$ and movable curve $v'\gamma_1\epsilon_1\big[(1-\tau)h - c^{-\eta}\big]^{\epsilon_1-1}$ as in figure 23. After thorough inspection, the non-linear equation is solvable at all the grids. Since we use the nonlinear solver, we must give some initial guess to every equation at each human capital value. An feasible guess for equation 49 is

$$c_{i,F}^n, c_{i,B}^n = h_i^{1/3}$$

The forward and backward approximations of discretized saving policy are calculated based on 49

$$\begin{aligned} s_{i,F}^n &= \gamma_1\big((1-\tau)h_i - c_{i,F}^n\big)^{\epsilon_1} + \gamma_2\big(\beta(\zeta)h_i^{-\zeta}\big)^{\epsilon_2} - \delta h_i + w \\ s_{i,B}^n &= \gamma_1\big((1-\tau)h_i - c_{i,B}^n\big)^{\epsilon_1} + \gamma_2\big(\beta(\zeta)h_i^{-\zeta}\big)^{\epsilon_2} - \delta h_i + w \\ e_{i,F}^n &= (1-\tau)h_i - c_{i,F}^n \\ e_{i,B}^n &= (1-\tau)h_i - c_{i,B}^n \end{aligned} \tag{50}$$

Initial Guess for Consumption and Value Function We need to guess the initial value when the interation number is $n = 0$ for all human capital grids $i = 1, 2, 3, ..., I-1$. Notice that the real value of consumption will not affect the equilibrium of consumption. All we need to consider is keeping the function form of initial value concave and increasing w.r.t human capital h. Since the CRRA utility function is concave and increasing, all we need to do is keep consumption function $c(h)$ monotonically increasing w.r.t h. Notice that the initial guess of consumption can not be too large to avoid zero savings at lower boundary. Moreover, Everyone lives on wages. If people have more endowments w, their savings will go up. When human capital is at the lower limit, consumption cannot be larger than wage, otherwise it will degrade directly. The power-education elasticity setting $e_t^{\epsilon_1} > 0$ in 3 keeps saving at lower boundary positive.

Here we divide the household's disposable income equally between the consumption function c_i^0 and education investment [1]

$$c_i^{0*} = \frac{(1-\tau)h_i}{2}, \quad (e_i^{0*})^{\epsilon_1} > 0$$

We guess the standardized CRRA form of value function

$$v_i^0 = \frac{u(c_i^0)}{\rho} = \frac{1}{\rho}\frac{(c_i^0)^{1-\eta} - 1}{1-\eta}$$

The forward and backward saving can be calculated through

$$\begin{aligned} s_{i,F}^0 &= c_i^0 - c_{i,F}^0 \\ s_{i,B}^0 &= c_i^0 - c_{i,B}^0 \end{aligned}$$

Let $(s_{i,F}^n)^+$ denote the positive part of forward approximation of saving, $(s_{i,B}^n)^-$ the negative part of backward one

$$\begin{aligned} (s_{i,B}^n)^- &= \min\{s_{i,B}^n, 0\} \\ (s_{i,F}^n)^+ &= \max\{s_{i,F}^n, 0\} \end{aligned}$$

We will update consumption current value according to forward or backward difference based on direction of saving. If saving is positive, the forward estimate of saving is subtracted from consumption, vise versa

$$c_i^1 = c_i^0 - (s_{i,B}^0)^- - (s_{i,F}^0)^+ \tag{51}$$

[1]Other form keeping $c^*(.)$ increasing w.r.t h is also feasible.

Then we update current value of saving and the present utility by

$$s_i^1 = \gamma_1\big((1-\tau)h_i - c_i^1\big)^{\epsilon_1} + \gamma_2(\beta(\zeta)h^{-\zeta})^{\epsilon_2} - \delta h_i + w$$
$$v_i^1 = \frac{(c_i^1)^{1-\eta}}{\rho(1-\eta)}$$

Given v_i^n, s_i^n and h_i^n, we can solve v_i^{n+1} through iteration matrix system. We do endless loop for $n = 1, 2, 3, ...$ according to the discrete HJB equation unless v_i^{n+1} is close enough to v_i^n or absolute value if their difference is smaller than same critical value *crit*.

A.1.2 Discretization Matrix System

Let i donates individual and I is the number of human capital grid points. Given v_i^n, to discretize the continous-time expression and v_i^{n+1} is now implicitly defined by the equation:

$$\frac{v_i^{n+1} - v_i^n}{\Delta} + \rho v_i^{n+1} = \frac{(c_i^n)^{1-\eta}}{1-\eta} + (v_i^{n+1})' s_i^n + \frac{\sigma^2}{2}(v_i^{n+1})'' h_i^2$$

$$s_i^n = \gamma_1\big((1-\tau)h_i - c_i^n\big)^{\epsilon_1} + \gamma_2(\beta(\zeta)h^{-\zeta})^{\epsilon_2} - \delta h_i + w$$

where

$$(v_i^{n+1})' s_i^n = (v_{i,F}^{n+1})'(s_{i,F}^n)^+ + (v_{i,B}^{n+1})'(s_{i,B}^n)^-$$
$$\big(v_i^{n+1}\big)'' = \frac{v_{i+1}^{n+1} - 2v_i^{n+1} + v_{i-1}^{n+1}}{(\Delta h)^2}$$

Given the upwind scheme first-derivative approximation and second-derivative approximation 48, the discretization of HJB equation is

$$\begin{aligned}\frac{v_i^{n+1} - v_i^n}{\Delta} + \rho v_i^{n+1} =& \frac{(c_i^n)^{1-\eta}}{1-\eta} + \frac{v_{i+1}^{n+1} - v_i^{n+1}}{\Delta h}(s_{i,F}^n)^+ + \frac{v_i^{n+1} - v_{i-1}^{n+1}}{\Delta h}(s_{i,B}^n)^- \\ &+ \frac{\sigma^2}{2}h_i^2 \frac{v_{i+1}^{n+1} - 2v_i^{n+1} + v_{i-1}^{n+1}}{(\Delta h)^2}\end{aligned}$$

Collecting terms with the same subscripts in the right-hand side and we end up with the discretized HJB equation

$$\frac{v_i^{n+1} - v_i^n}{\Delta} + \rho v_i^{n+1} = \frac{(c_i^n)^{1-\eta}}{1-\eta} + v_{i-1}^{n+1} x_i^n + v_i^{n+1} y_i^n + v_{i+1}^{n+1} z_i^n \tag{52}$$

$$
\begin{cases}
x_i^n = -\dfrac{(s_{i,B}^n)^-}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_i^2}{(\Delta h)^2} \\
y_i^n = -\dfrac{(s_{i,F}^n)^+}{\Delta h} + \dfrac{(s_{i,B}^n)^-}{\Delta h} - \sigma^2\dfrac{h_i^2}{(\Delta h)^2} \\
z_i^n = \dfrac{(s_{i,F}^n)^+}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_i^2}{(\Delta h)^2}
\end{cases}
\tag{53}
$$

Special Treatment for the Lower Boundary of Human Capital At the lower boundary $h = h_{\min} = h_1$, we make use of 58 and write the approximation as

$$
\begin{aligned}
\frac{v_1^{n+1} - v_1^n}{\Delta} + \rho v_1^{n+1} =& u(c_1^n) + \frac{v_2^{n+1} - v_1^{n+1}}{\Delta h}(s_{1,F}^n)^+ + \frac{v_1^{n+1} - v_0^{n+1}}{\Delta h}(s_{1,B}^n)^- \\
&+ \frac{\sigma^2}{2} h_1^2 \frac{v_2^{n+1} - 2v_1^{n+1} + v_0^{n+1}}{(\Delta h)^2}
\end{aligned}
\tag{54}
$$

What this equation says is, if we stand at the first wealth grid point h_1, $\frac{v_2^{n+1} - v_1^{n+1}}{\Delta h}(s_{1,F}^n)^+$ means when savings is positive, wealth increases into the second grid h_2, $v_2^{n+1} - v_1^{n+1}$ is the difference between two wealth grids. There's no constraint on borrowing, it's going to go forward, it's not going to go down, it doesn't matter in the calculation because $(s_{1,B}^n)^- = 0$.

Substitute $\left(v_{i,F}^n\right)'$ for $\left(v_{i,B}^n\right)'$, we have

$$
\begin{aligned}
\frac{v_1^{n+1} - v_1^n}{\Delta} + \rho v_1^{n+1} =& u(c_1^n) + \frac{v_2^{n+1} - v_1^{n+1}}{\Delta h}\left((s_{1,F}^n)^+ + (s_{1,B}^n)^-\right) \\
&+ \frac{\sigma^2}{2} h_1^2 \frac{v_3^{n+1} - 2v_2^{n+1} + v_1^{n+1}}{(\Delta h)^2}
\end{aligned}
$$

Collecting terms with the same subscripts in the right-hand side and we end up with the discretized HJB equation

$$
\frac{v_1^{n+1} - v_1^n}{\Delta} + \rho v_1^{n+1} = \frac{(c_1^n)^{1-\eta}}{1-\eta} + x_1 v_0^{n+1} + y_1^n v_1^{n+1} + z_1^n v_2^{n+1} + w_1 v_3^{n+1}
$$

$$
\begin{cases} x_1 = 0 \\ y_1^n = -\dfrac{(s_{1,F}^n)^+}{\Delta h} - \dfrac{(s_{1,B}^n)^-}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_1^2}{(\Delta h)^2} \\ z_1^n = \dfrac{(s_{1,F}^n)^+}{\Delta h} + \dfrac{(s_{1,B}^n)^-}{\Delta h} - \sigma^2\dfrac{h_1^2}{(\Delta h)^2} \\ w_1 = \dfrac{\sigma^2}{2}\dfrac{h_1^2}{(\Delta h)^2} \end{cases} \tag{55}
$$

or another simple expression:

$$
\begin{cases} x_1 = 0 \\ y_1^n = -\dfrac{(s_{1,F}^n)^+}{\Delta h} \\ z_1^n = \dfrac{(s_{1,F}^n)^+}{\Delta h} \\ w_1 = 0 \end{cases} \tag{56}
$$

Notice that actually $h_1 = 0$, thus the last term in y_1^n, z_1^n and w_1 can be drop out. And As in state boundary condition, we have $s(\underline{h}) > 0$ so that $(s_{1,B}^n)^- = 0$. For $w_1 = 0$, in short, if there were no assets there would be no risk volatility.

Special Treatment for the Upper Boundary of Human Capital For large h which referring to the people with highest income, we have $v(h) = \tilde{v}_0 + \tilde{v}_1 h^{1-\eta}, (\epsilon_1 = 1)$ or $v(h) = \tilde{v}_0 + \tilde{v}_1 h^{2-\eta-\epsilon_1}, (\epsilon_1 < 1)$ 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

$$
\begin{aligned} v''(h_I) &= -\eta\frac{v'(h_I)}{h_I}, \quad \epsilon_1 = 1 \\ v''(h_I) &= (1-\eta-\epsilon_1)\frac{v'(h_I)}{h_I}, \quad \epsilon_1 < 1 \end{aligned} \tag{57}
$$

From policy function and boundary condition, we have

$$
\frac{\sigma^2}{2} h_I^2 v''(h_I) = \xi v'(h_I) \tag{58}
$$

$$
\begin{aligned}
\xi &= -\eta\frac{\sigma^2}{2}h_I, \quad \epsilon_1 = 1 \\
\xi &= (1-\eta-\epsilon_1)\frac{\sigma^2}{2}h_I, \quad \epsilon_1 < 1
\end{aligned}
\tag{59}
$$

We will condition 58 below when solving HJB equation using a finite difference method. The state constraint 23 is equivalent to the situation when net wealth reaches a ceiling, consumption must be no less than saving, namely

$$
c(h_I) \geq s(h_I) = \gamma_1\big((1-\tau)h_I - c(h_I)\big)^{\epsilon_1} + \gamma_2(\beta(\zeta)h_I^{-\zeta})^{\epsilon_2} - \delta h_I + w
$$

At the upper boundary $h = h_{\max} = h_I$, we make use of 58 and write the approximation as

$$
\begin{aligned}
\frac{v_I^{n+1} - v_I^n}{\Delta} + \rho v_I^{n+1} =& u(c_I^n) + \frac{v_{I+1}^{n+1} - v_I^{n+1}}{\Delta h}(s_{I,F}^n)^+ + \frac{v_I^{n+1} - v_{I-1}^{n+1}}{\Delta h}(s_{I,B}^n)^- \\
&+ \frac{v_I^{n+1} - v_{I-1}^{n+1}}{\Delta h}\xi \\
\xi =& (1-\eta-\epsilon_1)\frac{\sigma^2}{2}h_I
\end{aligned}
\tag{60}
$$

We have $(s_{I,F}^n)^+ = 0$ and let $v_{I+1}^{n+1} - v_I^{n+1} = v_I^{n+1} - v_{I-1}^{n+1}$. The corresponding entries become:

$$
\begin{cases}
x_I^n = -\dfrac{(s_{I,F}^n)^+}{\Delta h} - \dfrac{(s_{I,B}^n)^-}{\Delta h} - \dfrac{\xi}{\Delta h} \\
y_I^n = \dfrac{(s_{I,F}^n)^+}{\Delta h} + \dfrac{(s_{I,B}^n)^-}{\Delta h} + \dfrac{\xi}{\Delta h} \\
z_I^n = \dfrac{(s_{I,F}^n)^+}{\Delta h}
\end{cases}
\tag{61}
$$

or another simple expression[1]

$$
\begin{cases}
x_I^n = -\dfrac{(s_{I,B}^n)^-}{\Delta h} - \dfrac{\xi}{\Delta h} \\
y_I^n = +\dfrac{(s_{I,B}^n)^-}{\Delta h} + \dfrac{\xi}{\Delta h} \\
z_I^n = 0
\end{cases}
\tag{62}
$$

[1]As in proposition 4.1, at the upper end of the state space, $s(h) < 0$ for $h > h_{\max}$ so that $(s_{I,F}^n)^+ = 0$.

Intensity Matrix Equation 52 with 53 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:

$$\frac{v^{n+1} - v^n}{\Delta} + \rho v^{n+1} = u^n + \mathbf{A_H}^n v^{n+1} \tag{63}$$

where

$$\mathbf{A_H}^n = \begin{bmatrix} y_1 & z_1 & w_1 & 0 & \cdots & 0 \\ x_2 & y_2 & z_2 & 0 & \ddots & \vdots \\ 0 & x_3 & y_3 & z_3 & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 & x_I & y_I \end{bmatrix}_{I \times I}, \quad u^n = \begin{bmatrix} u\left(c_1^n\right) \\ \vdots \\ \vdots \\ \vdots \\ u\left(c_I^n\right) \end{bmatrix} \tag{64}$$

Let

$$\mathbf{B}^n = \left(\frac{1}{\Delta} + \rho\right) \mathbf{I} - \mathbf{A_H}^n \tag{65}$$

$$\mathbf{b}^n = \mathbf{u}^n + \frac{1}{\Delta} \mathbf{v}^n \tag{66}$$

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.[1] In particular, all rows sum to zero, diagonal elements are non-positive and off-diagoal elements are non-negative.

The discretized HJB system can be written as the linear system. We solve it using sparse matrix routines.

$$\mathbf{B}^n \mathbf{v}^{n+1} = \mathbf{b}^n \tag{67}$$

Let $\Delta \to \infty$ (when programming, we set a relative large number, for example, 1000), 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

$$\rho \mathbf{v}^{n+1} = \mathbf{u}^n + \mathbf{A_H}^n \mathbf{v}^{n+1} \tag{68}$$

[1] In HACT, various types of stochastic processes are included. The finite difference method basically approximates this process such as with a discrete Poisson process with a transition matrix $\mathbf{A}^n$ summarizing the corresponding Poisson intensities while satisfing all the properties Poisson transition matrix needs to satisfy.

Now from the numerical solution to HJB equations we can solve the sequence of stationary consumption, saving and value function. [2]

A.2 Solving KF equation: Implicit Method using MFG System (FD Method)

A.2.1 Upwind Scheme

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

$$\begin{aligned}&-\frac{(s_{i,F})^+ f_i - (s_{i-1,F})^+ f_{i-1}}{\Delta h} - \frac{(s_{i+1,B})^- f_{i+1} - (s_{i,B})^- f_i}{\Delta h} \\ &+ \frac{\sigma^2}{2} \frac{h_{i+1}^2 f_{i+1} - 2h_i^2 f_i + h_{i-1}^2 f_{i-1}}{(\Delta h)^2} = 0\end{aligned} \tag{69}$$

Collecting terms, we have

$$f_{i-1} z_{i-1} + f_i y_i + f_{i+1} x_{i+1} = 0 \tag{70}$$

$$\begin{cases} x_{i+1} = -\dfrac{\left(s_{i+1,B}^n\right)^-}{\Delta h} + \dfrac{\sigma^2}{2} \dfrac{h_{i+1}^2}{(\Delta h)^2} \\ y_i = -\dfrac{\left(s_{i,F}^n\right)^+}{\Delta h} + \dfrac{\left(s_{i,B}^n\right)^-}{\Delta h} - \sigma^2 \dfrac{h_i^2}{(\Delta h)^2} \\ z_{i-1} = \dfrac{\left(s_{i-1,F}^n\right)^+}{\Delta h} + \dfrac{\sigma^2}{2} \dfrac{h_{i-1}^2}{(\Delta h^2)} \end{cases}$$

or a more concise way of writing

[2]In an instance of I data, they are stored in a form as $\mathbf{v}_{1\times I} = [v_1, v_2, v_3, \cdots, v_{I-1}, v_I]$, $\mathbf{c}_{1\times I} = [c_1, c_2, c_3, \cdots, c_{I-1}, c_I]$, $\mathbf{s}_{1\times I} = [s_1, s_2, s_3, \cdots, s_{I-1}, s_I]$.

$$
\begin{cases}
x_i = -\dfrac{\left(s_{i,B}^n\right)^-}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_i^2}{(\Delta h)^2} \\
y_i = -\dfrac{\left(s_{i,F}^n\right)^+}{\Delta h} + \dfrac{\left(s_{i,B}^n\right)^-}{\Delta h} - \sigma^2 \dfrac{h_i^2}{(\Delta h)^2} \\
z_i = \dfrac{\left(s_{i,F}^n\right)^+}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_i^2}{(\Delta h)^2}
\end{cases}
\tag{71}
$$

A.2.2 Discretization Matrix System

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. [1] 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.

$$
-\frac{\left(s_{1,F}\right)^+ f_1}{\Delta h} + \frac{\sigma^2}{2}\frac{h^2 f_2}{(\Delta h)^2} = 0
$$

Thus,

$$
f_0 z_0 + f_1 y_1 + f_2 x_2 = 0
$$

The special treatment for the lower boundary reads

$$
\begin{cases}
z_0 = 0 \\
y_1 = -\dfrac{\left(s_{1,F}\right)^+}{\Delta h} \\
x_2 = \dfrac{\sigma^2 h_2^2}{2(\Delta h)^2}
\end{cases}
\tag{72}
$$

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 :

$$
f_{I-1} z_{I-1} + f_I y_I + f_{I+1} x_{I+1} = 0
$$

[1] However with low income shock in HACT, here are no drifting out of this boundary, for there is a mass point.

The special treatment for the upper boundary reads

$$
\begin{cases}
\tilde{x}_I = x_I = -\dfrac{\left(s_{I,B}^n\right)^-}{\Delta h} + \dfrac{\sigma^2}{2}\dfrac{h_I^2}{(\Delta h)^2} \\
\tilde{y}_I = y_I + z_I = \dfrac{\left(s_{I,B}^n\right)^-}{\Delta h} - \dfrac{\sigma^2}{2}\dfrac{h_I^2}{(\Delta h)^2} \\
\tilde{z}_I = \dfrac{\sigma^2}{2}\dfrac{h_I^2}{(\Delta h)^2}
\end{cases}
\tag{73}
$$

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 69 with 71 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

$$
\mathbf{0} = \mathbf{A_K}^{\mathrm{T}}(\mathbf{v})\mathbf{f} \tag{74}
$$

$$
\mathbf{A_k} = \begin{bmatrix}
y_1 & x_1 & 0 & 0 & \cdots & 0 \\
z_2 & y_2 & x_2 & 0 & \ddots & \cdots \\
0 & z_3 & y_3 & x_3 & \ddots & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots & \ddots \\
\vdots & \ddots & \ddots & z_{I-2} & y_{I-2} & x_I \\
0 & \ddots & \ddots & 0 & z_{I-1} & y_I
\end{bmatrix} \tag{75}
$$

A.3 Another Way to Solve KF Equation (Iterative Method)

As described in the paper, we develop some iterative algorithms.

We have gotten the numerical solution of $c(h)$ from HJB numerical solutions part. In this part, we can solve the human capital distribution $f(h)$ sequence throught the KF equation. Now turn to the solution of stochastic optimal control Kolmogorov Forward Equations satisfying the property of probability density :

Intergrating the KF equation 18, we get

$$
\frac{\sigma^2}{2}\frac{\partial}{\partial h}[h^2 f(h)] = s(h)f(h) + C_0
$$

where C_0 is a constant.

Writting equation in the first-order linear differential equation form as

$$\frac{df(h)}{dh} + p(h)f(h) = q(h)$$

where $\begin{cases} p(h) = 2[\frac{1}{h} - \frac{s(h)}{\sigma^2 h^2}] \\ q(h) = \frac{2C_0}{\sigma^2 h^2} \end{cases}$

An integrating factor $\mu(h)$ can be defined as $\mu(h) = e^{\int p(h)dh} = e^{2\int(\frac{1}{h} - \frac{\phi(h)}{\sigma^2 h^2})dh}$. Multiply the first-order linear equation by $\mu(h)$, we get

$$\mu(h)\frac{df(h)}{dh} + p(h)\mu(h)f(h) = \mu(h)q(h)$$

Integrate the last equation, we get the general solution. Let $C_2 = \frac{2C_0}{\sigma^2}$ and the general solution is $f(h) = C_1 f_1(h) + C_2 f_2(h)$, where two particular solutions satisfying

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

It can be written as a variable upper bound integral form as

$$\begin{aligned} f(h) &= C_1 f_1(h) + C_2 f_2(h) \\ &= C_1 e^{-2\int_{h0}^h p(x)dx} + +C_2\int_{h_0}^h x^{-2}e^{-2\int_x^h p(z)dz}dx \end{aligned} \tag{76}$$

where boundary conditions are the same as in equations 21 and 22.

A.3.1 Interative Method (1)

Interative 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 13 and 76 in the two boundary conditions, we have the equation system

$$
\begin{aligned}
C_1\int_{\underline{h}}^{\infty} e^{-\int_{h_0}^{h} p(x)dx}dh + C_2\int_{\underline{h}}^{\infty}\int_{h_0}^{h} x^{-2}e^{-\int_{x}^{h} p(z)dz}dxdh = 1 \\
C_1\int_{\underline{h}}^{\infty} he^{-\int_{h_0}^{h} p(x)dx}dh + C_2\int_{\underline{h}}^{\infty} h\int_{h_0}^{h} x^{-2}e^{-\int_{x}^{h} p(z)dz}dxdh = \int_{\underline{h}}^{\infty} hf(h)dh
\end{aligned}
\tag{77}
$$

where $H(\tau) = \int_{\underline{h}}^{\infty} hf(h)dh$.

Recall the notes 41

$$
\begin{cases}
g(h) = e^{-\int_{h_0}^{h} p(x)dx} = e^{-\int_{h_0}^{h} \frac{2}{x} - \frac{2s(x)}{\sigma^2 x^2}dx} = (h^{-2} - h_0^{-2})e^{\int_{h_0}^{h} \frac{2s(x)}{\sigma^2 x^2}dx} \\
z(x) = x^{-2}e^{\int_{h_0}^{x} p(z)dz} = x^{-2}e^{\int_{h_0}^{x} \frac{2}{z} - \frac{2s(z)}{\sigma^2 z^2}dz} \\
G = \int_0^{\infty} g(h)dh \\
G_h = \int_0^{\infty} hg(h)dh \\
Z = \int_0^{\infty} g(h)\int_1^{h} z(x)dxdh \\
Z_h = \int_0^{\infty} hg(h)\int_1^{h} z(x)dxdh
\end{cases}
$$

The equation system 77 can be written as

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

From iterative method (1) we prove that Z and Z_h 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 77. C_1 is calculated by the expression

$$
C_1 = \frac{1}{G} = \frac{1}{\int_{\underline{h}}^{\infty} e^{-2\int_{h_0}^{h} \frac{1}{x} - \frac{s(x)}{\sigma^2 x^2}dx}dh} = \frac{1}{\int_{\underline{h}}^{\infty}(h^{-2} - h_0^{-2})e^{2\int_{h_0}^{h} \frac{s(x)}{\sigma^2 x^2}dx}dh}
$$

After sloving two constants, we have $f(h_0) = C_1$. Notice that $f'(h) = \frac{f(h+\Delta h) - f(h)}{\Delta h}$, and we can iterate the equation from $f(h_0)$ as

$$
\frac{f(h+\Delta h) - f(h)}{\Delta h} + 2\Big[\frac{1}{h} - \frac{s(h)}{\sigma^2 h^2}\Big]f(h) = \frac{C_2}{h^2} = 0
$$

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

In discretization form, 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.

$$C_1' = \frac{\int_{\underline{h}}^{\infty} hf(h)dh}{\int_{\underline{h}}^{\infty} hg(h)dh} = \frac{H(f(h))}{G_h} \tag{78}$$

If $C_1 \sim C_1'$, then the equation system is satisfied.

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 $f(1) = C$, still use the iteration equation

$$\frac{f(h + \Delta h) - f(h)}{\Delta h} + 2\big[\frac{1}{h} - \frac{s(h)}{\sigma^2 h^2}\big] f(h) = 0$$

s.t.

$$\sum_{i=1}^{I-1} \frac{f(h_i) + f(h_{i+1})}{2} dh = 1$$

$$\sum_{i=1}^{I-1} \frac{s(h_i)f(h_i) + s(h_{i+1})f(h_{i+1})}{2} dh = 0$$

All we need to do is normalize the density aggregation to 1.

Check the Consistency 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

$$||f - g|| = \sup_{h \in S} |f(h) - g(h)|, \quad S = [\underline{h}, \bar{h}]$$

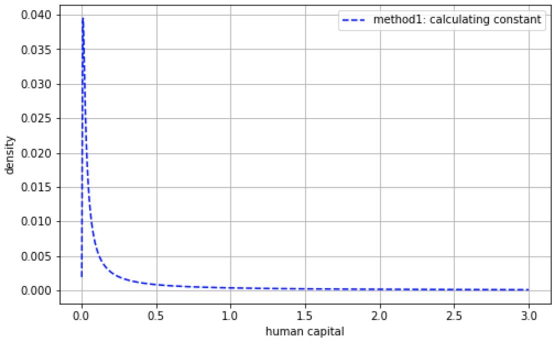


Figure 24: Method1

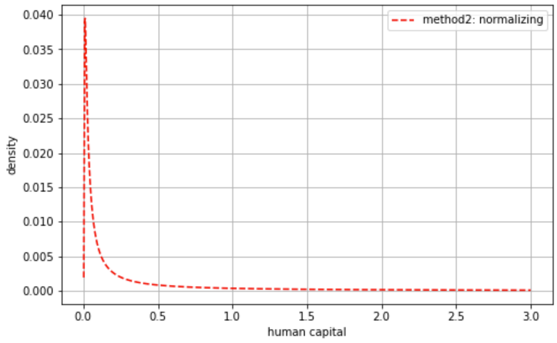


Figure 25: Method2

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 Bisection Method to Solve the Public Education Payments under Balanced Government Budget

To solve the government public education expenditure in state-independent tax-transfer model 9 and education subsidy model 12. We supplement the numerical method for solving MFG system in HACT(2020) with an algorithm for solving tax-expenditure equilibrium. In the process of numerical solution of HJB equation, an arbitrarily constant β is given to solve the household's policy functions which are put into the numerical KF equation program. After obtaining density of human capital, we can easily calculate the total social tax $\tilde{T}$. According to the condition of government budget balance, $\tilde{T}$ is regarded the transfer payment in equilibrium.

$$T = \int_{\underline{h}}^{\infty} \left(\pi + \beta(\zeta)h^{-\zeta}\right) f(h)dh$$

$$\tilde{T} = \tau \int_{\underline{h}}^{\infty} hf(h)dh$$

We solve out the β^* at equilibrium

$$\beta^*(\zeta) = \tau \frac{\int_{\underline{h}}^{\infty} hf(h)dh - \pi}{\int_{\underline{h}}^{\infty} h^{-\zeta} f(h)dh}$$

Assuming continuous function $g(T) = T - \tilde{T}$ in the interval $[T_{\min}, T_{\max}]$, we want to solve $g(T) = 0$. The intervals for positive tax and negative tax are different, but the procedure is the same. Given the maximum number of loops m, the convergence standard

$crit$, the upper and lower limits of the interval is updated according to the size relationship between $T$and $\tilde{T}$.

$$\begin{cases} T_{\min}^{n+1} = T^n, \tilde{T}^n - T^n \geq crit \\ T_{\max}^{n+1} = T^n, T^n - \tilde{T}^n \leq crit \end{cases} \tag{79}$$

where $n = 0, 1, 2, ..., m$ is loop count.

$$T^{n+1} = \frac{T_{\min}^{n+1} + T_{\max}^{n+1}}{2}$$

Do it repeatedly until $|T^n - \tilde{T}^n| < crit$ or $n \geq m$. Let $\beta = \beta^*$, and we get the income distribution at equilibrium.

C Proof of Proposition 4.1: Behaviour for the High-Income Households

Let's note the general form of value function $v(h) = Bh^a$ where a and B are constants.

Take the education subsidy model as example. Recall the HJB equation 14

$$\rho v(h) = \max_{c(h)} u(c(h)) + v'(h)\left[w + \gamma_1((1-\tau)h - c(h))^{\epsilon_1} + \gamma_2\left(\pi + \beta h^{-\zeta}\right)^{\epsilon_2} - \delta h\right] + \frac{\sigma^2 h^2}{2} v''(h)$$

Write it with the Hamiltonian function $H(v'(h)) = \max_{c(h)}\{u(c(h)) - v'(h)\gamma_1\left[(1-\tau)h - c(h)\right]^{\epsilon_1}\}$,

$$\rho v(h) = \max_{c(h)} H(v'(h)) + v'(h)\left[\gamma_2 w + \left(\pi + \beta h^{-\zeta}\right)^{\epsilon_2} - \delta h\right] + \frac{\sigma^2 h^2}{2} v''(h)$$

Let $\frac{dH(v'(h))}{dc(h)} = 0$ and we solve the first-order condition equation 17

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

C.1 If $\epsilon_1 = 1$

Hamiltonian function reads $H(v'(h)) = \max_{c(h)}\{u(c(h)) - v'(h)\gamma_1 c(h)$

$$
\begin{aligned}
s(h) =& w + \big(\gamma_1(1-\tau) - \delta\big)h + H'(v'(h)) + \gamma_2\left(\pi + \beta h^{-\zeta}\right)^{\epsilon_2} \\
=& w + \big(\gamma_1(1-\tau) - \delta\big)h - \gamma_1 c(h) + \gamma_2\left(\pi + \beta h^{-\zeta}\right)^{\epsilon_2}
\end{aligned}
$$

Let $\frac{dH(v'(h))}{dc(h)} = 0$ and F.O.C. solves

$$
u'(c) = \gamma_1 v'(h) \tag{80}
$$

Using $c(h)$ and their derivatives to express $v'(h), v''(h)$ and $v'''(h)$,

$$
\begin{cases}
v'(h) = & \frac{1}{\gamma_1}u'(c) = \frac{1}{\gamma_1}c(h)^{-\eta} \\
v''(h) = & \frac{1}{\gamma_1}u''(c)c'(h) = -\frac{\eta}{\gamma_1}c(h)^{-\eta-1}c'(h) \\
v'''(h) = & \frac{1}{\gamma_1}u'''(c)c'(h)^2 + \frac{1}{\gamma_1}u''(c)c''(h) = \frac{\eta(\eta+1)}{\gamma_1}c(h)^{-\eta-2}c'(h)^2 - \frac{\eta}{\gamma_1}c(h)^{-\eta-1}c''(h)
\end{cases}
$$

Substitute them into the Euler equation in Lemma 4.1

$$
\begin{aligned}
\left[\frac{\rho+\delta}{\gamma_1} - (1-\tau)\right]u'(c) = & \frac{1}{\gamma_1}u''(c)c'(h)s(h) + \frac{1}{\gamma_1}\frac{\sigma^2 h^2}{2}\left[u'''(c)\left(c'(h)\right)^2 + u''(c)c''(h)\right] \\
& + \frac{1}{\gamma_1}\sigma^2 h u''(c)c'(h)
\end{aligned} \tag{81}
$$

Let's divide both sides of the Euler equation 81 by $c(h)u''(h)$,

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

Notice that under the "asymptotical linearity consumption" assumption $\frac{c(h)}{h} \sim \phi$, we have $\lim\limits_{h\to\infty} c'(h) = \phi$, $\lim\limits_{h\to\infty} c''(h) = 0$ and $\lim\limits_{h\to\infty} \frac{c'(h)h}{c(h)} = 1$. The relative risk aversion (RRA) $\eta = -\frac{c(h)u''(c)}{u'(c)}$ is bounded above for all c. [1] Thus we have

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

$$\frac{s(h)}{h} \sim \vartheta = -\frac{\rho+\delta}{\eta} + \gamma_1 \frac{(1-\tau)}{\eta} - (1-\eta)\frac{\sigma^2}{2} \tag{82}$$

So far we have proved this property by solving F.O.C. Now if us start from the general form of Euler equation rather than soving out the exact F.O.C. equation 80, let's see if we can get the same conclustion. The proof is as follows.

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

$$\vartheta = -\frac{\rho+\delta}{\eta} + \frac{\gamma_1\big(1-\tau\big)}{\eta} - (1-\eta)\frac{\sigma^2}{2} \tag{83}$$

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

$$\begin{aligned}\phi =& 1-\tau-\frac{\vartheta+\delta}{\gamma_1}\\ =&\frac{(1-\tau)(\eta-1)}{\eta} + \frac{(1-\eta)\frac{\sigma^2}{2}-\delta}{\gamma_1} + \frac{\rho+\delta}{\gamma_1\eta}\end{aligned} \tag{84}$$

Given value function, we could get the same conclusion as in equation 82. The results of the two methods are consistent and the constant coefficient of value function solves

$$B = \frac{\phi^{-\eta}}{\gamma_1(1-\eta)} \tag{85}$$

C.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.

$$u'(c) = v'(h)\gamma_1\epsilon_1 \left[(1-\tau)h - c(h)\right]^{\epsilon_1 - 1}$$

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

Divide the both sides of the first-order condition in equation 17 with $h^{-\eta}$ to homogenize it

$$\left(\frac{c(h)}{h}\right)^{-\eta} = (2-\eta-\epsilon_1)B\gamma_1\epsilon_1 \left[1-\tau-\frac{c(h)}{h}\right]^{\epsilon_1-1} \tag{86}$$

Then the asymptotic-ratio constant ϕ in 30 solves

$$\phi^{-\eta} = (2-\eta-\epsilon_1)B\gamma_1\epsilon_1[1-\tau-\phi]^{\epsilon_1-1}$$

See lemma 4.1, we get the derivative of HJB (Euler equation) by dividing both sides on h of this formula

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

Recalling $\frac{c(h)}{h} \sim \phi$, as $h \to \infty$ we have [1]

$$s'(h) \sim \gamma_1\epsilon_1\big(1-\tau-\phi\big)^{\epsilon_1} h^{\epsilon_1-1} - \delta \tag{87}$$

Since $0 < \epsilon_1 < 1$, the derivative of saving as $h \to \infty$ is simplified as [2]

$$s'(h) \sim -\delta \tag{88}$$

Euler equation

$$\begin{aligned}\rho aBh^{a-1} =& aBh^{a-1}\gamma_1\epsilon_1\big[(1-\tau)h - c(h)\big]^{\epsilon_1-1}\phi + aBh^{a-1}s'(h) \\ &+ (a-1)aBh^{a-2}\big(s(h)+\sigma^2 h\big) + (a-2)(a-1)aBh^{a-3}\frac{\sigma^2 h^2}{2}\end{aligned}$$

[1] Recall $s(h) = w + \gamma_1((1-\tau)h - c(h))^{\epsilon_1} + \gamma_2 T^{\epsilon_2} - \delta h$. The derivative of saving solves $s'(h) = \gamma_1\epsilon_1 h^{\epsilon_1-1}\big(1-\tau-\frac{c(h)}{h}\big)^{\epsilon_1-1}(1-\tau-c'(h)) - \delta$.

[2] Here $v'(h) = aBh^{a-1}$, $v''(h) = (a-1)aBh^{a-2}$, $v'''(h) = (a-2)(a-1)aBh^{a-3}$.

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

$$\rho = -\delta + (a-1)\vartheta + a(a-1)\frac{\sigma^2}{2} \tag{89}$$

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

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

$$\vartheta = \gamma_1 h^{\epsilon_1 - 1}((1-\tau)h^{1-\epsilon_1} - \phi)^{\epsilon_1} - \delta$$

$$\vartheta = -\frac{\rho + \delta}{1-a} - \frac{\sigma^2 a}{2} = \frac{\rho + \delta}{1-\eta-\epsilon_1} - \frac{2-\eta-\epsilon_1}{2}\sigma^2 \tag{90}$$

$$\phi = 1 - \tau - \frac{((\vartheta + \delta)h - w - \gamma_2 T^{\epsilon_2})^{\frac{1}{\epsilon_1}}}{h\gamma_2^{\frac{1}{\epsilon_1}}} \tag{91}$$

Since $s'(h) = -\delta$ and $\frac{s(h)}{h} \sim \vartheta$, if

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

$$\phi = 1 - \tau + \frac{(w + \gamma_2 T^{\epsilon_2})^{\frac{1}{\epsilon_1}}}{h\gamma_2^{\frac{1}{\epsilon_1}}}$$

$$\phi \sim 1 - \tau$$

D Proof of Proposition 4.3: the Inverse Gamma Distribution Pattern

Recall KE equation written as a variable upper bound integral form as the first-order differential equation of variable upper bound integral form in proposition 4.3.

$$f(h) = C_1 e^{-2\int_{\underline{h}}^{h} p(x)dx} + C_2 \int_{\underline{h}}^{h} x^{-2} e^{-2\int_x^h p(z)dz} dx$$

$$\begin{cases} p(h) = 2[\frac{1}{h} - \frac{s(h)}{\sigma^2 h^2}] \\ q(h) = \frac{2C_0}{\sigma^2 h^2} \end{cases}$$

and the notes we give in equations 41

$$\begin{aligned}
&g(h) = e^{-\int_{h_0}^{h} p(x)dx} = e^{-\int_{h_0}^{h} \frac{2}{x} - \frac{2s(x)}{\sigma^2 x^2} dx} = (h^{-2} - h_0^{-2}) e^{\int_{h_0}^{h} \frac{2s(x)}{\sigma^2 x^2} dx} \\
&z(x) = x^{-2} e^{\int_{h_0}^{x} p(z)dz} = x^{-2} e^{\int_{h_0}^{x} \frac{2}{z} - \frac{2s(z)}{\sigma^2 z^2} dz} \\
&G = \int_0^\infty g(h)dh \\
&G_h = \int_0^\infty hg(h)dh \\
&Z = \int_0^\infty g(h) \int_1^h z(x)dxdh \\
&Z_h = \int_0^\infty hg(h) \int_1^h z(x)dxdh
\end{aligned}$$

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

$$s(h) \sim \vartheta h$$

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

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

As $h- > \infty$,

$$\begin{aligned}
g(h) &= h^{-2} e^{2\int_1^h \frac{s(x)}{\sigma^2 x^2}} dx \sim h^{-2} e^{2\int_1^h \frac{\vartheta}{\sigma^2 x}} dx \\
&= h^{-2 + \frac{2\vartheta}{\sigma^2}}
\end{aligned}$$

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

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

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

Denote $\kappa = \frac{2\vartheta}{\sigma^2}$ and integrate functions $g(h)$ and $hg(h)$ respectively

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

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

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.

$$z(x) = x^{-2}e^{2\int_1^x \frac{1}{z} - \frac{s(z)}{\sigma^2 z^2}dz} \sim x^{-\frac{2\kappa}{\sigma^2}}$$

Similarly, as $x- > \infty$, integral values solve

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

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

Thus as h goes to infinity, Z and Z_h goes to infinity .

Recall the equations

$$1 = C_1 G + C_2 Z$$

$$H = C_1 G_h + C_2 Z_h$$

According to KF equation 18, $C_2 = \frac{2C_0}{\sigma^2}$ 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 the result in proposition 4.3,

$$f(h) = Ce^{-\frac{\beta}{h}} h^{-\alpha-1} \tag{92}$$

whose integration constant

$$C = \frac{1}{\int_0^\infty e^{-2\int_{h_0}^h \frac{1}{x} - \frac{s(x)}{\sigma^2 x^2} dx} dh}$$

The human capital in this setting indeed follows Inverse Gamma distribution.

E Proof of Proposition 4.2: Pareto Tail of Human Capital Distribution

Intergrating KF equation 18, we get

$$\frac{\sigma^2}{2} \frac{\partial}{\partial h} [h^2 f(h)] = s(h) f(h) + C_0 \tag{93}$$

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

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

Now we define $y(h) = \frac{\sigma^2 h^2 f(h)}{2}$, and rewrite equation 93 as

$$\begin{aligned} y'(h) &= \frac{2s(h)}{\sigma^2 h^2} y(h) = 2\frac{\tilde{s}+\bar{s}h}{\sigma^2 h^2} y(h) \\ &= \mu_1 \frac{y(h)}{h} + \mu_2(h) y(h) \end{aligned} \tag{94}$$

where we define

$$\mu_1 = \frac{2\bar{s}}{\sigma^2}$$
$$\mu_2(h) = \frac{2\tilde{s}}{\sigma^2 h^2} = 2\frac{\tilde{s}+\bar{s}h}{\sigma^2 h^2} - 2\frac{\bar{s}}{\sigma^2 h}$$

Dividing equation 94 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.

$$\lim_{h\to\infty} \ln(\frac{y(h)}{h^{\mu_1}}) = \bar{\epsilon} \tag{95}$$

Recall the definition $y(h) = \frac{\sigma^2 h^2 f(h)}{2}$, we have

$$\lim_{h\to\infty} \frac{\sigma^2 h^2 f(h)}{2h^{\mu_1}} = \exp(\bar{\epsilon})$$

$$\lim_{h\to\infty} \frac{f(h)}{h^{\mu_1}} = \frac{2\exp(\bar{\epsilon})}{\sigma^2 h^2}$$

$$\lim_{h\to\infty} \frac{f(h)}{h^{\mu_1-2}} = \frac{2\exp(\bar{\epsilon})}{\sigma^2}$$

To sum up, we claim when $h \to \infty$,

$$f(h) \sim \epsilon h^{-\zeta-1}$$
$$\epsilon = \frac{2\bar{\epsilon}}{\sigma^2}$$
$$\zeta = 1 - \mu_1 = 1 - \frac{2\bar{s}}{\sigma^2}$$

F Estimate the Quantile of Income Distribution

Given quantile q, corresponding numerical human capital bound follows the quantile function $F^{-1}(.)$

$$F^{-1}(q) = \sup\{h|F(h) \leq q\}, \quad q \in [0,1] \tag{96}$$

where the finite and positive first moment is defined as

$$E(h) = \int_0^1 F^{-1}(q)dq = \int_0^\infty hf(h)dh. \tag{97}$$

Any income distribution supported on the nonnegative halfline with a finite and positive first moment admits a Lorenz curve, representing the share of total income received by u fraction of the lower-income households. The continuous version of Lorenz curve $L(u)$ can be written as

$$L(u) = \frac{\int_0^u F^{-1}(q)dq}{E(h)}, \quad u \in [0,1]. \tag{98}$$

or

$$L(u) = \frac{\int_0^{h_u} hf(h)dh}{\int_0^\infty hf(h)dh}, \quad u \in [0,1]. \tag{99}$$

where h_u is the human capital value at the share of total income u.

In discrete case of I-points household income, Lorenz curve value $L_k(I)$ (or accumulated income share) represent the share of total income received by the $u = \frac{k}{I} \times 100\%$ of the lower-income households. Define the discrete version of accumulated income share $L_k(h)$, $k = 0, 1, 2, ..., I$.

$$L_k(I) = \frac{\sum_{i=1}^k h_i f_i(h)dh}{\sum_{i=1}^I h_i f_i(h)dh} \tag{100}$$

Then,

$$L_k(I) = L_{k-1}(I) + \frac{h_k f_k(h)dh}{\sum_{i=1}^{I} h_i f_i(h)dh}, \quad k = 2, ..., I. \tag{101}$$

$$L_0(I) = 0$$

$$L_1(I) = \frac{h_1 f_1(h)dh}{\sum_{i=1}^{I} h_i f_i(h)dh}$$

When we estimate the quantile of income distribution, we will encounter the probability of a certain h human capital value just across the loci such as 20. At this point, since the h value at this loci is equal, it is evenly divided. Part of the segmentation makes the cumulative probability equal to 0.2, and the other part is divided to the next quantile segment. The wealth (human capital) proportion in total human capital between two quantiles $p1$ and $p2$ can be calculated as

$$\pi(p1, p2) = L_{p2}(h) - L_{p1}(h)$$

.

One of the numerous definitions of the Gini index is given by as twice the area between the Lorenz curve L and the "equality line" $L(h) = h$,

$$Gini(h, f(h)) = 1 - 2\int_0^{\infty} L(h)f(h)dh \tag{102}$$

Thus the discrete form Gini coefficient is

$$Gini(h, f(h)) = 1 - 2\sum_{i=1}^{I} \frac{L_{i-1}(h) + L_i(h)}{2} f_i(h)dh \tag{103}$$

Moreover, Aitchison and Brown (1957) [1] has shown that we can derive Gini coefficient from a normal distribution [1] and it can be generalized to any known distribution. Since we've proved that human capital follows an inverse Gamma distribution. Applying it to human capital h, the Gini-coefficient of the inverse Gamma distribution [2] is

[1] Aitchison and Brown (1957) show that the Gini-coefficient is given by: $Gini(X) = 2\Phi(\frac{\sigma_h}{\sqrt{2}} - 1)$, where $\Phi(.)$ is the CDF of normal distribution. Using a first-Taylor approximate we get $Gini \approx \frac{\sigma_h}{\sqrt{\pi}}$.

[2] It is computed in McDonald (1984).

$$Gini(h) = \frac{\Gamma(\alpha - 0.5)}{\Gamma(\alpha)\sqrt{\pi}}$$

which decreases with the value of α with the threshold case of $Gini(\alpha = 1) = 1$.