Public Debt and Wealth Inequality

By Xiaolu Li and Yang Tang and Jingjing Zhu and Shenghao Zhu*

The issuance of public debt affects the return on assets in the market, and then affects the macroeconomic equilibrium and wealth distribution. We use a two-period overlapping generations model with idiosyncratic investment risk to analyze the impact of public debts on Macroeconomy. Using a perturbation method, we find the the influence results of the disturbance problem. We use welfare maximization to investigate the optimal debt. *Keywords: Public debt; Two-period OLG model; Welfare; Pertubation*

The level and type of government debt has long been an important fiscal policy concern (Aiyagari and McGrattan (1998)). The issuance of public debt provides liquidity to finance the market and can affect market asset yields , which can and in turn affects macroeconomic equilibrium and wealth distribution, and improve risk sharing by increasing liquidity in the economy (Flodén (2001)). In this paper, we develop a two-period overlapping generations model with idiosyncratic investment risk to study the optimal size of public debts.

Due to the ravages of the new coronavirus, governments in recent years have pursued aggressive debt financing policies to stimulate the economy. Until 2022, the Federal Reserve announced a 75 basis point interest rate hike at the end of July, which was the fourth rate hike so far this year, and its largest concentrated rate hike since the early 1980s. The latest news is that the Associated Press reported on October 4 that the total public debt of the United States has exceeded $31 trillion, and the debt amount is getting closer to the statutory maximum limit of about

* Xiaolu Li: Nanjing University of Posts and Telecommunications, xiaoluleez@gmail.com. Yang Tang: Nanyang Technological University, tangyang@ntu.edu.sg. Jingjing Zhu: University of International Business and Economics, University of International Business and Economics, 10 Huixin Dongjie, Beijing, China, email: zhujingjing0121@163.com. Shenghao Zhu: School of International Trade and Economics, University of International Business and Economics, 10 Huixin Dongjie, Beijing, China, email: zhushenghao@yahoo.com.

$31.4 trillion set by Congress on the borrowing capacity of the U.S. government.

In response to this phenomenon, the Peterson Foundation noted that U.S. fiscal spending was already on an unsustainable path prior to the COVID-19 outbreak, and the outbreak rapidly exacerbated the U.S. fiscal challenges. Against the backdrop of the U.S. government's continued expansion of fiscal spending and the Federal Reserve's increasing interest rate hikes, major neighboring Latin American economies are facing serious challenges such as rising inflation and local currency depreciation, and the "brutal growth" of U.S. government debt has raised concerns from all walks of life.

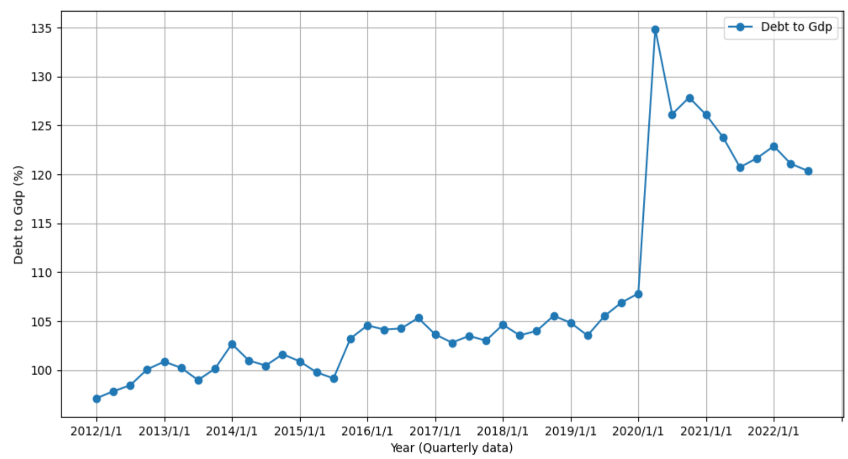

Figure 1. Total Public Debt as Percent of Gross Domestic Product (2012/1/1-2022/7/1)

Figure 1, we used the Federal Reserve's data [1], which reports the ratio of U.S. public debt to GDP in recent years. We can see that the ratio rose sharply around 2020 and then fell, fluctuating about 125% for the rest of the year. The economic fallout of the COVID-19 health crisis is likely to accelerate these patterns going forward, as governments pursue aggressive debt-financed stimulus policies.

Based on this background, we explore whether we can find the optimal size of the public debt issuance. What is the relationship between welfare and inequality? We

[1]Web site address: https://fred.stlouisfed.org/series/GFDEGDQ188S

also use a flow-tracking approach to find the impact of the public debt on different income groups of the population.

We develop a two-period overlapping-generations dynamic macroeconomic model with heterogeneous households to study the optimal size of the public debt. Households face idiosyncratic income shocks, including labor efficiency shocks and the investment return risk. The economy has no aggregate risk. There is a continuum of measure 1 families in the economy. Each family consists of one parent and one child. Individual is young in the first period of his life and becomes old in the second period. Each old agent gives birth to one child. The population of the economy keeps constant. When agents are in the young period, they receive wages as workers and inheritance. Spending includes consumption, investment and the purchase of public debt. The income sources of agents in their old-age include profits from operating an enterprise, capital incomes from previous investment and interest on public debt. Expenses include consumption and bequests left to children.

To evaluate the quantitative significance of our findings, we begin with a detailed calibration of the model that replicates the US economy, including the distribution of wealth and earnings. Thus, in the model, the top 1% appear to match the data well, at least with respect to their key economic characteristics. We then use the calibrated version of the model to quantitatively determine the answers to the above questions.

We study the impacts of public debts on the wealth distribution. Public debts influence debt returns and private firm returns in the general equilibrium. Thus public debts influence the wealth distribution. We investigate the impacts of public debts on the tail of the wealth distribution. The stationary distribution is the fixed point of a functional equation, which describes law of motion of the wealth distribution through the transition probability. To find the stationary distribution, we use machine learning to solve the functional equation. We can easily construct the loss function using the functional equation. We show that the individual wealth distribution has a unique stationary distribution with an asymptotic Pareto tail.

We also use the perturbation method to find the impact of public debts on individuals with different income levels. Clarify the basis for determining the most public debt level through the perturbation decomposition of the welfare distribution. This decomposition is particularly useful in considering the impact of changes in the public debt because it allows us to measure the impact on each of the components.

The main marginal contributions of the article are the following three points. First we explore the relationship between public debt and wealth inequality using a two-stage generational overlap model. Second, the idea that uninsurable labor income risks may justify the optimal public debt is not new but we find stable distributions using a machine learning approach to decompose the impact of national debt into channels. Finally, based on the parameters of the model and the social welfare maximization principle, we find the optimal level of national debt.

I. Literature review

Modeled as early as in Barro (1979) develops a simple theory of 'optimal' public finance that identifies some factors that would influence the choice between taxes and debt issue. We conduct our analysis with the help of an incomplete markets Aiyagari-Bewley framework , which Aiyagari (1994) combines Brock and Mirman (1972)'s theory on the standard growth model and Bewley (1986)'s theory about idiosyncratic labor endowment shocks, provide an exposition of models whose aggregate behavior is the result of market interaction among a large number of agents subject to idiosyncratic shocks.

One the one hand, the Aiyagari-Bewley framework is widely used in the study of various types of fiscal policies. We divide existing works into three groups based on their methodologies.

(1) NEOCLASSICAL GROWTH MODEL. — Several papers analyze fiscal problem using a representative agent neoclassical growth model and emergence of a body of literature directly analyze the government debt budget constraint, the seminal work of Aiyagari and McGrattan (1998), Flodén (2001), and Nakajima and Takahashi (2017), among others.

In this economy households are exposed to idiosyncratic income shocks but no aggregate risk. These shocks are uninsurable because insurance markets are absent. Aiyagari and McGrattan (1998) using the heterogeneous agent trading in risk-free assets capital and government debt, with endogenous labor supply incomplete market model and data from the United States after World War II find that the welfare gains at the optimal level are negligible worries about high debt levels in the United States or other economies are misplaced. Under the similar framework, Flodén (2001) argued that both debt and transfers can significantly improve risk sharing and that transfers would be more effective than public debt if a utilitarian welfare criterion is used. When the government is allowed to choose transfers optimally, the role of public debt disappears and the optimal level of debt is -100% of output. Nakajima and Takahashi (2017) extend this models by incorporating consumption taxes, and calibrate the Japanese economy for the 1995–2013 period, analyze the influence of huge government debt on social welfare and find that the optimal level of government debt is –50% of GDP for Japan.

(2) Ramsey problem. — In other literature to study the optimal Ramsey plan for a broad set of fiscal instruments in an environment with incomplete markets, such as Acikgoz (2015), Gottardi, Kajii and Nakajima (2015), Röhrs and Winter (2017) and Acikgoz et al. (2018), etc. Especially Dyrda and Pedroni (2021) consider complete-markets versions of model in which can analytically characterize the optimal fiscal policy, the Ramsey problem in this quantitative general equilibrium model of heterogeneous agents and uninsurable special risks is named the Standard Incomplete Market (SIM) model.

Aiyagari (1995) finds that the tax on capital must be positive at the steady-state solution of the Ramsey problem. But Gottardi, Kajii and Nakajima (2015) in addition to considering government bonds and physical capital, human capital is added, show that the existence of a steady state imposes no real restriction on the value of the tax rate on capital and the optimality of a positive tax rate is primarily determined by the comparison of costs and benefits of the tax and debt.

Again, Acikgoz (2015) concludes in Aiyagari (1995) this premise is premature in an infinite-horizon model with incomplete markets and heterogeneous agents, and he Use the necessary conditions for optimization shown that long-term income tax rates, levels of government debt, and the distribution of wealth and consumption can be studied independently of the transition path and without taking a position on the initial conditions of the economy. Acikgoz et al. (2018) prove that there is a unique long-term optimal level of government debt that is independent of the initial debt level, the optimal long-term level of government debt is 1.1 times GDP.

(3) Overlapping generations model. — Several papers use an overlapping generations model. Wan and Zhu (2019) use a decomposition technique to investigate the impact of estate taxes on the long-run wealth inequality not the optimal issuance of national debt, and find that the different results of estate taxes are due to the different redistribution effects.

One the other hand, in the last 40 years labor earnings, market income and wealth inequality have increased substantially in the U.S. at the top end of the distribution (Kindermann and Krueger (2014)). Based on Aiyagari-Bewley framework, some studies use a persistent skewed distribution of stochastic earnings to explain the wealth distribution and wealth inequality, such as Castaneda, Diaz-Gimenez and Rios-Rull (2003), De Nardi (2004), Alonso-Ortiz and Rogerson (2010), Kindermann and Krueger (2014), Chang (2022), etc.

De Nardi (2004) examining the impact of bequests on welfare by establishing intergenerational relationships between two generations of legacy additions and find that voluntary bequests can explain the emergence of large estates, which are often accumulated in more than one generation, and characterize the upper tail of the wealth distribution in the data. The relationship between marginal tax rates and welfare for the top 1% part of the income distribution is mainly studied in Kindermann and Krueger (2014). Chang (2022) propose a different perspective about the study of the links between central bank policy and inequality, he do not offer an answer to the question of what the social objective function should be, but the question of what the central bank's mandate should be. And he point that in the presence of heterogeneity and the central bank lacks commitment power, the central bank mandate should be less egalitarian than the social welfare function. Similarly, Röhrs and Winter (2017) find that if inequality is large, the optimal level of debt that maximizes the steady-state welfare is even lower and should be negative, -0.8. One reason why the optimal level of debt is low or even negative when steady-state welfare is maximized is that this criterion ignores the welfare loss of reducing debt along the transition path to a low-debt steady state.

Commentary. — Our paper builds on previous studies in the literature and builds on the heterogeneous agents and idiosyncratic income shocks model by internalizing payroll tax rates and externalizing estate and capital taxes to investigate not only the optimal amount of public debt issuance, but also the corresponding welfare and inequality relationships using a two-stage OLG model.

The remainder of the paper proceeds as follows. The benchmark model is introduced including environment and general equilibrium in section II. Section III the distribution of wealth and young-age income, the social welfare and the equation for the perturbation of public debt are introduced respectively. Section IV, we use the percentile results of the wealth distribution and young-age income distribution in the model to match the actual situation in the United States, and determine the parameters used in this paper. Section V and Section VI report the numerical and perturbation results of the model respectively. Finally, we concluding remarks are contained in section VII.

II. The benchmark Model

We set up a two-period overlapping-generations model with idiosyncratic risk. There is a continuum of measure 1 families in the economy. Each family consists of one parent and one child. Individual is young in the first period of his life and becomes old in the second period. Each old agent gives birth to one child. The population of the economy keeps constant. We describe the balanced growth path of such an economy without insurance markets but with trading in risk-free assets (debt) and risk assets (capital). The young agent work as workers and are inelastic in providing 1 unit labor. He becomes an entrepreneur and operates a firm when he is old. The financial market is incomplete and borrowing is not allowed. We also assume that there are no aggregate shocks.

A. Environment

Government issues debt in each period to pay them back in the form of taxes. There are three forms of taxation: payroll tax (τ_w), corporate tax (τ_p) and inheritance tax (τ_z), the inheritance income after-tax is $\overline{z} = (1 - \tau_z)z$. In particular, $\overline{w} = (1 - \tau_w)w$ represents the after-tax wage rate. We give the rules of conduct for government in the market. The supply of public debt B_t is assumed to be exogenous. The government decides the payroll tax τ_w to balance its budget in each period, R is the gross interest rate and endogenously determined,

$$R_t B_t = B_{t+1} + \tau_{wt} w_t + \tau_{zt} \int z_t + \tau_{pt} \int \pi_{t+1}(\theta_{t+1}, k_{t+1}).$$

Individual have two stages of consumption. He can choose to spend in the first stage (when he is young, c_t) or in the second stage (when he is old, c_{t+1}) to spend and leave part of his inheritance (z_{t+1}) to his children. Specifically, individual preference is featured with "joy-of-giving" bequest motive. While they receive bequest from their parents, they also leave bequest to their children to gain utilities.

The agent derives utility in period t from consumption, this utility is given by $c_{1,t}^{1-\gamma}/(1-\gamma)$; and from consumption and inheritance in period $t+1$, this utility is given by $c_{2,t+1}^{1-\gamma}/(1-\gamma) + \chi\left[(1-\tau_z)z_{t+1}\right]^{1-\gamma}/(1-\gamma)$, where γ is the relative risk aversion, χ measures the weights of bequest in the preferences, β is the subjective discount factor. We also denote $\rho_t = c_{1,t}/y_t$ to be the expenditure share on consumption in the young-age, the remaining expenditure then go to investment. Denote $\phi_{t+1} = k_{t+1}/(k_{t+1} + b_{t+1})$ to be the fraction of capital in the investment portfolio, the remaining are then invested in the public debt, where k_{t+1} and b_{t+1} denote the holding of capital and public debt in old-age, respectively.

Definition II.1 *The optimal policy functions of individuals ($c_{1,t}$, ρ_t, ϕ_{t+1}, $c_{2,t+1}$, z_{t+1}) are given by*

1. *($c_{2,t+1}$, z_{t+1}) solves*
$$\max_{c_{2,t+1};z_{t+1}} \frac{c_{2,t+1}^{1-\gamma}}{1-\gamma} + \chi \frac{[(1-\tau_{z,t+1})z_{t+1}]^{1-\gamma}}{1-\gamma}, \ s.t. \quad c_{2,t+1} + z_{t+1} = h_{t+1};$$
2. *($c_{1,t}$, ρ_t, ϕ_{t+1}) solves*
$$\max_{c_{1,t}} \frac{c_{1,t}^{1-\gamma}}{1-\gamma} + \beta E\left(\frac{c_{2,t+1}^{1-\gamma}}{1-\gamma} + \chi \frac{[(1-\tau_{z,t+1})z_{t+1}]^{1-\gamma}}{1-\gamma}\right), \ s.t. \quad c_{1,t} + k_{t+1} + b_{t+1} = y_t;$$
3. *and,* $y_t = (1-\tau_{zt})z_t + (1-\tau_{wt})\eta_t w_t$,
$$h_{t+1} = (1-\tau_{p,t+1})\pi_{t+1}(\theta_{t+1}, k_{t+1}) + (1-\delta)k_{t+1} + R_{t+1}b_{t+1}.$$

where y_t and h_{t+1} are incomes when the young-age and when the old-age. There is a continuum of mass 1 of workers indexed by their labor shock $\eta \in \Upsilon$, $\log(\eta) \sim N(\mu_\eta, \sigma_\eta^2)$ and the actual mean of η, $exp(\mu_\eta + \sigma_\eta^2/2)$, must be equal to 1.

The old agent creates a firm, hires young workers to produce the unique type of consumption goods available in the economy. Production requires both capital and labor. Specifically, the production function takes form:

$$Y = \theta A k^\alpha l^{1-\alpha}, \alpha \in (0,1),$$

Definition II.2 *We assume competitive product and factor markets, the optimal firm problem is (l_{t+1}, π_{t+1}) such that*

1. *(l_{t+1}, π_{t+1}) solves* $\max\limits_{l_{t+1}} \theta_{t+1} A k_{t+1}^\alpha l_{t+1}^{1-\alpha} - w_{t+1} l_{t+1}$

where the aggregate technology level is denoted as A, w_{t+1} is the wage rate. And θ is an idiosyncratic productivity shock.

The idiosyncratic productivity is assumed to follow a truncated logarithmic normal distribution with a mean of μ_θ and a standard deviation of σ_θ, Θ is the set of all θ_i. And, we have: $\theta \in \Theta$, where $log(\theta) \sim N(\mu_\theta, \sigma_\theta^2)$.

Proposition II.1 *The consumption ($c_{2,t+1}$) of the elderly, the inheritance retention (z_{t+1}), the labor demand (l_{t+1}) and the profit (π_{t+1}) are then given by*

$$l_{t+1} = \left[\frac{\theta_{t+1}A(1-\alpha)}{w_{t+1}}\right]^{\frac{1}{\alpha}} k_{t+1}, \tag{1}$$

$$\pi_{t+1}(\theta_{t+1}, k_{t+1}) = \alpha(\theta_{t+1}A)^{\frac{1}{\alpha}} \left(\frac{w_{t+1}}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}} k_{t+1}, \tag{2}$$

$$c_{2,t+1} = h_{t+1}\frac{1}{1+\chi^{\frac{1}{\gamma}}(1-\tau_{z,t+1})^{\frac{1-\gamma}{\gamma}}}, \tag{3}$$

$$z_{t+1} = \frac{h_{t+1}\varphi}{(1-\tau_{z,t+1})}, \varphi = \frac{\chi^{\frac{1}{\gamma}}(1-\tau_{z,t+1})^{\frac{1}{\gamma}}}{1+\chi^{\frac{1}{\gamma}}(1-\tau_{z,t+1})^{\frac{1-\gamma}{\gamma}}}. \tag{4}$$

The inheritance income after-tax is $\overline{z}_{t+1} = h_{t+1}\varphi$, thus φ is the after-tax inheritance income ratio, represents the proportion of income left to children in old-age

Proposition II.2 *The optimal policy functions of individuals are then given by*

$$c_{1,t} = \rho_t[(1-\tau_{zt})z_t + (1-\tau_{wt})\eta_t w_t], \tag{5}$$

$$k_{t+1} = \phi_{t+1}(1-\rho_t)[(1-\tau_{zt})z_t + (1-\tau_{wt})\eta_t w_t], \tag{6}$$

$$E\left\{[\omega_{t+1}(\theta_{t+1})\phi_{t+1} + R_{t+1}(1-\phi_{t+1})]^{-\gamma}(\omega_{t+1}(\theta_{t+1}) - R_{t+1})\right\} = 0, \tag{7}$$

$$\rho_t = \frac{\left\{E[\omega_{t+1}(\theta_{t+1})\phi_{t+1} + R_{t+1}(1-\phi_{t+1})]^{1-\gamma}\right\}^{-\frac{1}{\gamma}}}{(\beta\Gamma)^{\frac{1}{\gamma}} + \left\{E[\omega_{t+1}(\theta_{t+1})\phi_{t+1} + R_{t+1}(1-\phi_{t+1})]^{1-\gamma}\right\}^{-\frac{1}{\gamma}}}. \tag{8}$$

Where, $\omega_{t+1}(\theta_{t+1}) = (1-\tau_{p,t+1})\alpha\left(\theta_{t+1}A\right)^{\frac{1}{\alpha}}\left(\frac{w_{t+1}}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}} + 1 - \delta.$

We use $J(\theta) = \omega(\theta)\phi + R(1-\phi)$ for return on capital, describes the return on assets of 1 dollar under the optimal portfolio choice, where ϕ represents the proportion of investment in physical capital, $(1-\phi)$ represents the proportion of investment in debt. Ultimately, $\Gamma = \left(1 + \chi^{\frac{1}{\gamma}}\left(1-\tau_{z,t+1}\right)^{\frac{1-\gamma}{\gamma}}\right)^{\gamma}$, the optimization problem of a household is given by the following indirect utility functional equation, u:

$$u(y_t) = \frac{y_t^{1-\gamma}}{1-\gamma}\left\{\rho_t^{1-\gamma} + \beta\Gamma(1-\rho_t)^{1-\gamma}E\left\{J(\theta)^{1-\gamma}\right\}\right\}. \tag{9}$$

Proof. See Appendix A.A1 for more details.

B. General equilibrium

We consider the players in the market to form an equilibrium market outcome. The equilibrium liquidation results of labor market and public debt market are given.

Definition II.3 *Given the exogenous variable the supply of public debt (B_t), technology level (A) , and taxation rate $\{\tau_{pt}, \tau_{zt}\}$, the equilibrium consists a series of quantities $\{\ c_{1,t}, c_{2,t+1}, k_t, b_t\ \}$,a sequence of prices $\{w_t, R_t\}$ and payroll tax (τ_{wt}) such that the following holds:*

1. *Given prices and tax levels, the individual maximizes utility (Definition II.1).*
2. *Given prices and tax levels, entrepreneurs maximize profits (Definition II.2).*
3. *The government decides the payroll tax τ_w to balance its budget in each period.*
4. *The market has cleared (labor market, public debt market).*

Labor market clear, (w_{t+1}) solves $\int l_{t+1} = 1$.

Public debt markets clear, (R_t) solves $\int b_t = B_t$, and $R_t B_t = B_{t+1} + \tau_{wt} w_t + \tau_{zt} \int z_t + \tau_{pt} \int \pi_{t+1}(\theta_{t+1}, k_{t+1})$.

Proposition II.3 *In the steady state, the exogenous variable (w, K, R) are given by*

1. *(w) Given the equilibrium conditions of the labor market,*

$$w = (1-\alpha) A K^{\alpha} \left(E \theta^{\frac{1}{\alpha}} \right)^{\alpha}. \tag{10}$$

2. *(K) The aggregate economy has a unique steady state, capital is jointly determined by the risk-free interest rate R, the consumption ratio ρ, the investment ratio ϕ and the tax rate τ_w,*

$$K^{\alpha-1} = \frac{\dfrac{1}{\phi(1-\rho)} - \varphi(1-\delta) - \dfrac{(1-\phi)\varphi R}{\phi}}{A\left(E\theta_t^{\frac{1}{\alpha}} \right)^{\alpha} \{\varphi\alpha(1-\tau_p) + (1-\tau_w)(1-\alpha)\}}. \tag{11}$$

3. *(R) Combine the government decides and the public debt market clear, R is determined by the consumption ratio ρ, the investment ratio ϕ and the tax rate τ_w,*

$$R = \frac{1 + \dfrac{\left[\dfrac{\left[(1-\alpha)\tau_w + \tau_p\alpha + \dfrac{\tau_z\varphi}{1-\tau_z}\alpha(1-\tau_p)\right]}{(1-\phi)(1-\rho)} + \dfrac{(1-\delta)\varphi\phi}{1-\phi}\left(\dfrac{\tau_z}{1-\tau_z}(1-\tau_w)(1-\alpha) - (1-\alpha)\tau_w - \tau_p\alpha \right) \right]}{\varphi\alpha(1-\tau_p) + (1-\tau_w)(1-\alpha)}}{1 + \dfrac{\varphi}{\varphi\alpha(1-\tau_p) + (1-\tau_w)(1-\alpha)} \left[(1-\alpha)\tau_w + \tau_p\alpha - \dfrac{\tau_z}{1-\tau_z}(1-\alpha)(1-\tau_w) \right]}. \tag{12}$$

Proof. *See Appendix A.A2 for the above.*

According to the analysis of Proposition II.2, the expressions of the endogenous variables are ultimately determined by three important variables: the consumption share ρ, the fraction of capital ϕ and the payroll tax rate τ_w.

III. Distribution and welfare

According to Díaz-Giménez, Glover and Ríos-Rull (2011), *wealth* as the net worth of households, and define $x_{t+1} = k_{t+1} + b_{t+1}$ as individual's wealth. Novel techniques are now available to identify and characterize the impact of monetary policy on distributional outcomes. We use probability density functions to give distributions of wealth $w(x)$ and the distribution of young-age income $g(y)$, and machine learning applications are used to obtain a stable distribution, where the distribution results and loss functions are presented in the Appendix B.B4.

Distribution. — The dynamics of wealth (x_t) and young-age income (y_t) can be summarized as follows,

$$x_t: \quad x_{t+1} = (1-\rho_t)\left\{J(\theta)x_t\varphi + (1-\tau_{wt})w_t\eta_t\right\},$$

$$y_t: \quad y_{t+1} = y_t(1-\rho_t)J(\theta)\varphi + (1-\tau_{w,t+1})\eta_{t+1}w.$$

Let $w_{X_t}(x)$ be the probability density function of x_t, the cumulative distribution function of x_{t+1}, where $\nu = \dfrac{x-(1-\rho)\overline{w}\eta}{(1-\rho)\varphi J(\theta)}$, is:

$$W_{t+1}(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_t(\nu)p(\theta)q\left(\frac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\frac{(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}d\theta d\nu. \tag{13}$$

Let $g_{Y_t}(y)$ be the probability density function of y_t. the cumulative distribution function of y_{t+1}, where $\mu = \dfrac{y-\eta\overline{w}}{(1-\rho)\varphi J(\theta)}$, is:

$$G_{t+1}(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_t(\mu)p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{(1-\rho)\varphi J(\theta)}{\overline{w}}d\theta d\mu. \tag{14}$$

Proposition III.1 *The individual wealth distribution has a unique stationary distribution with an asymptotic Pareto tail of an exponent κ, i.e.*

$$\lim_{x\to\infty}\frac{1-W(x)}{x^{-\kappa}} = C,$$

with $C > 0$. And κ solves

$$E[(1-\rho)\varphi J(\theta)]^{\kappa} = 1.$$

Definition III.1 (The stationary distribution) *For all $x \geq 0$ and $y \geq 0$, the stationary distribution of $\{x_t\}_{t=0}^{\infty}$ and $\{y_t\}_{t=0}^{\infty}$,*

1. *$w(x)$ solves $w(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(\nu)\, p(\theta) q\left(\frac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\frac{1}{(1-\rho)\overline{w}} d\theta d\nu.$*
2. *$g(y)$ solves $g(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\mu)\, p(\theta) q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{\overline{w}} d\theta d\mu.$*

Proof. *See Appendix B.B1 and B.B2 for all above.*

Social welfare. — The indirect utility function measures individual welfare levels, we aggregate the Equation (9) above to obtain the aggregate welfare about the young-age income probability density function $g(y)$,

$$V(y) = \int u(y)g(y)dy = \frac{\rho^{1-\gamma} + \beta\Gamma(1-\rho)^{1-\gamma}E\left\{J(\theta)^{1-\gamma}\right\}}{1-\gamma}\int y^{1-\gamma}g(y)dy. \tag{15}$$

Proof. *See Appendix B.B3.*

IV. Calibration

We calibrate the initial stationary equilibrium of the model to replicate key properties relevant for the shape of the optimal public debt and welfare in the U.S. economy. Each period in the model corresponds to 20 year.

Calibration from literature. — To calibration the model, we first fix a few parameters to the levels typically assumed in the literature (see Table 1). There is no uniform standard for the selection of these parameters in the article, we determined the values of the required parameters based on previous literature. The literature is documented as follows:

(β) the subject discount factor with 1-year is set to be 0.9511 (Gottardi, Kajii and Nakajima (2015)), 0.96 (Aiyagari (1994)), 0.98 (Acikgoz (2015)) and from 0.9806 to 1 (Aiyagari and McGrattan (1998)).

TABLE 1—CALIBRATION FROM LITERATURE

Para.	Description	Value
$\beta^{(20)}$	subject discount factor	0.98^{20}
$\delta^{(20)}$	capital depreciation rate	$1-(1-0.075)^{20}$
α	capital share	0.30
γ	risk aversion coefficient	2

(δ) the depreciation rate of capital with 1-year is set to be 0.06 (Gottardi, Kajii and Nakajima (2015)), 0.07 (Röhrs and Winter (2017)), 0.075 (Aiyagari and McGrattan (1998), Flodén (2001)), 0.08 (Aiyagari (1994), Acikgoz (2015)). In Acikgoz et al. (2018) calibrate the model with 10-year periods, adjust the 10-year depreciation rate to 0.720.

(γ) the risk aversion coefficient such as Attanasio et al. (1999), Gourinchas and Parker (2002) and De Nardi (2004), the value of the risk aversion factor is usually set in the range of [1,3], especially in Flodén (2001), Acikgoz (2015) and Acikgoz et al. (2018) set to be 2.

(α) the share of income that goes to capital, it is generally set between 0.30 and 0.40, such as 0.30 (Aiyagari and McGrattan (1998),Röhrs and Winter (2017)), 0.36 (Prescott (1986), Gottardi, Kajii and Nakajima (2015)and Acikgoz et al. (2018)) and 0.38 (Dyrda and Pedroni (2021)).

CALIBRATION FROM US DATA. — To calibrate the model, we set some parameters within the controllable range according to the actual data in the United States (see Table 2).

(τ_p) since July 1987, American corporate income tax has been divided into three levels of excess progressive tax of 15%, 25% and 34%. This paper sets the tax rate target at the lowest level of 15%.

(τ_z) inheritance tax rates, according to US policy, range from 18% to 40%, choose the lowest standard in our article.

(χ) we fit a linear relationship between personal income and consumption, and the marginal propensity coefficient of consumption was obtained at about 0.725. The

Table 2—Calibration from US data

Para.	Description	Match	Source/Target
τ_p	profit tax	0.15	15%, 25% and 34%
τ_z	bequest tax	0.18	$\tau_z \in [18\%, 40\%]$
χ	bequest motive coefficient	10.00	Propensity to consume =0.22
A	TFP	1.00	TFP $\in$ [0.75,1.02]

Note: Income:*https://fred.stlouisfed.org/series/PI*; Consumption:*https://fred.stlouisfed.org/series/PCEC*; TFP:*https://fred.stlouisfed.org/series/RTFPNAUSA632NRUG*

corresponding marginal propensity coefficient of save is around 0.27. Inheritance motivation should be considered in old-age, therefore we determine the value of χ according to the bequest motive and set consumption propensity below 0.27 in old age.

(A) the total factor productivity is the constant in the production function, which is normalized De Nardi (2004).

Benchmark Calibration. — The benchmark parameter value are summarized in Table 3. We internalize the tax rate τ_w to determine the tax rate in different cases based on the fact that the supply of bonds equals the demand in the market. The range of τ_w is specified to be (0,1). The idiosyncratic productivity and labor shock

Table 3—Benchmark Calibration

Para.	Description	Value
τ_w	payroll tax	$\in$ (0,1)
μ_θ	Log mean of prod disturbance	0.211
σ^2_θ	Log variance of prod disturbance	0.70
μ_η	Log mean of labor shock	-1
σ^2_η	Log variance of labor shock	2

are assumed to follow a truncated log-normal distribution, it is truncated to avoid the concern that the inequality is driven by the fat tail of the distribution.

Calibration results:. — We will report the stable distribution of wealth found using the probability density fixed point method. We compare them with U.S.

data, the numbers of data of wealth distribution in Table 4 and income distribution in Table 5 from Díaz-Giménez, Glover and Ríos-Rull (2011), they calculated the data from the Survey of Consumer Finances (SCF).

We use percentile tables of wealth and income to show that the parameters selected in the paper are consistent with actual results in the United States. According to the analysis in Equation (13), the stable distribution of wealth is reported in Table 4.

Table 4—The Wealth Distribution

	Wealth Partition							
Percentile	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
W_share_data	-0.002	0.011	0.045	0.112	0.120	0.111	0.267	0.336
W_share_model	0.061	0.038	0.053	0.096	0.105	0.111	0.251	0.285

Note: The first row are the wealth shares we computed from using the real data, the second row are their model counterparts.

Income is divided into labor income (wages) and capital income. Therefore, we define *income* as wage income when young-age and firm profit income when old-age.

Table 5—The Income Distribution

	Income Partition							
Percentile	0-20	20-40	40-60	60-80	80-90	90-95	95-99	99-100
Inc_share_data	0.028	0.067	0.113	0.183	0.138	0.102	0.159	0.210
Inc_share_model	0	0.001	0.011	0.028	0.052	0.173	0.494	0.241

Note: The first row are the income shares we computed from using the data, the second row are their data counterparts.

The wealth and income of our calibrated model match the target moments reasonably well. Top 1% groups share in the wealth and income distribution fits well. Therefore, we determined the parameter setting of the model according to the distribution results.

V. Numerical Exercises

Through calibration, we obtain a set of parameters that match the actual situation in the United States. Using this set of parameters, we then search for the optimal case of the public debt. The aim is to quantify the role of public debt in shaping the distribution of young-age income and wealth. In previous literature, the method of random number sampling is mostly used for the numerical solution of debt, such as Aiyagari and McGrattan (1998). The difference in this article is the use of machine learning to solve stable distributions.

Endogenous variable solution. — The main focus of this paper is to quantify the role of public debt. We choose to endogenize the wage tax rate, i.e., $\tau_w \in (0, 1)$, use the supply and demand balance of debt in the market to determine tax rates at different levels, and report the resulting equilibrium results, i.e., the numerical solutions of all endogenous variables (solve Equation (7), (8), (11) and (12) see Table D1).

Due to the increase in supply, the market interest rate (R) decreases. There is a negative correlation between public debt and tax rate. Rising debt and falling interest rates have characterized advanced economies over the past 40 years (Mian, Straub and Sufi (2021)). We validate the relationship between high debt and low interest rate, which Mian, Straub and Sufi (2021) develops a new framework to tackle the relationship between high debt and low interest rate. The framework shows how rising income inequality and the liberalization of the financial sector can push economies into a low interest rate-high debt environment.

From the point of view of the government, the increase of public debt is conducive to the government to pull the market domestic demand, which helps stimulate investment, therefore the corresponding capital investment (K), capital ratio (ϕ), and wage rate (w) all increase with the increase of public debt. From the perspective of firms and individuals, the lower tax rate increases the amount of money in the hands of individuals, at the same time, it will crowd out a part of individuals consumption,

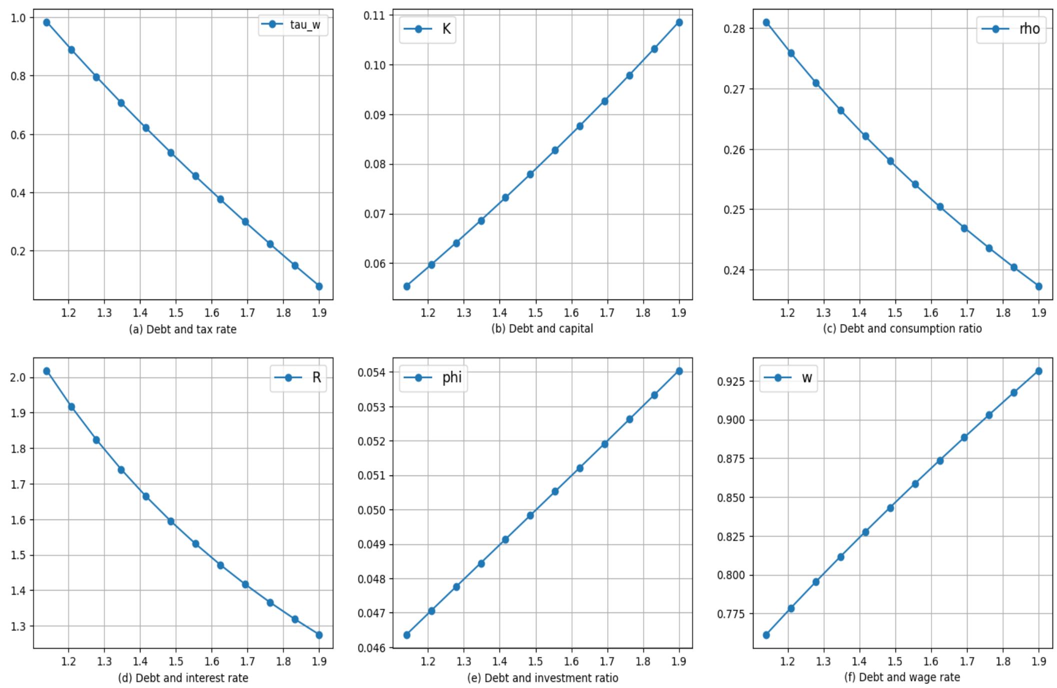

FIGURE 2. THE RELATIONSHIP BETWEEN ENDOGENOUS VARIABLES AND PUBLIC DEBT

reduces the ratio of consumption to income (ρ).

Figure 3 matches that of America's public debt in recent years and the possible future debt-to-GDP ratio is projected. The relationship between debt level and GDP in Figure 3(a), shows that both have the same growth trend. But the public debt is growing faster than GDP (See Figure 3(b)).

WELFARE. — According to the stable distribution of young-age income and the Equation (15), the relationship diagram of welfare and debt is obtained, as shown in Figure 4. With the increase of debt, debt to GDP showed a state of continuous increase, the highest increase to 143%, when welfare is optimal, the debt-to-GDP ratio reaches 1.267.

For the stable distribution solution of income, we use a machine learning approach to search. The process of machine learning is to construct the loss function from the target data and the predicted data. In our paper we are using the common

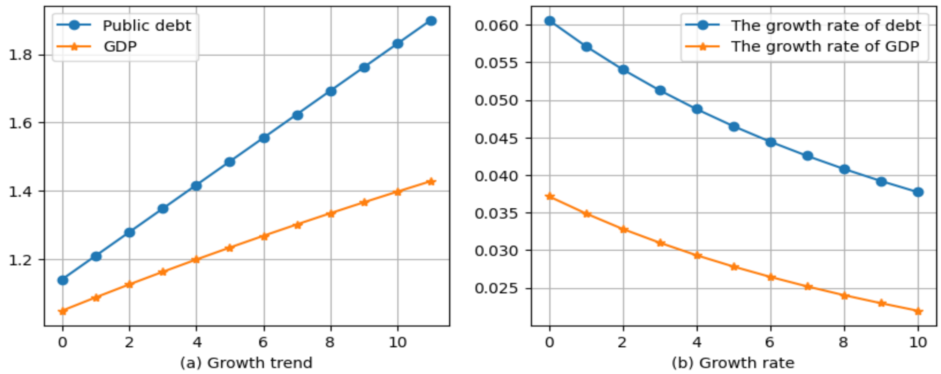

FIGURE 3. GDP AND PUBLIC DEBT

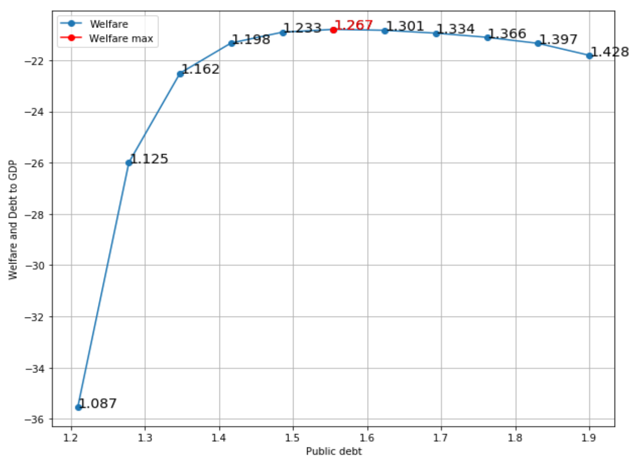

FIGURE 4. WELFARE AND PUBLIC DEBT

Note: The number of each point is the value of Debt to GDP.

mean square error (MSE) to continuously find the optimal combination of $g(y)$ to minimize the loss. The search process is a two-part process involving forward and backward propagation. Forward propagation is similar to a blind man going down a hill, standing on the hill and looking for the way down, given a random initial point of $g(y)$, to find the fastest way forward. Backward propagation is to feed the result of each loss back to the $g(y)$ that was guessed at the beginning for a new weight estimation and a new forward propagation process. From the perspective of the central bank mandate, the utilitarian welfare maximization problem is due to its incentive effect in favor of society and not from the perspective of social inequality, thus this outcome is best even if the outcome is unequal (Chang (2022)).

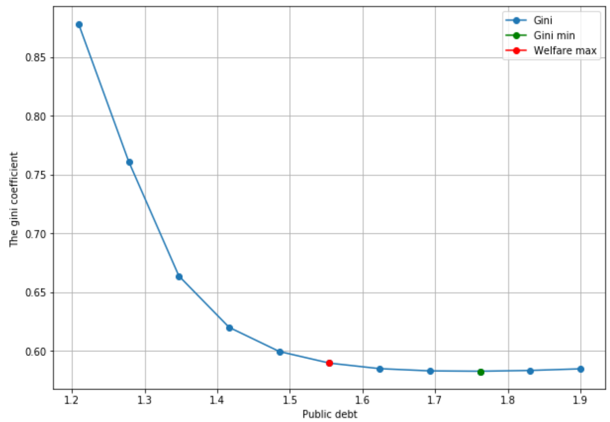

Figure 5. The Gini coefficient and public debt

Figure 5 reports the change of Gini coefficient, with the increase of public debt, Gini coefficient decreases first and then increases. When welfare is at its maximum, inequality is already small, but still not at its most equal. Our consideration of this phenomenon is that there has been no precise definition of the relationship between welfare and inequality, and we think it may be because social welfare aims to enhance the well-being of all people by the state or the government through some means. According to Maslow's Hierarchy of Needs, the rich mostly cross the physiological needs and security needs, while the poor stay at the lowest level, it is very difficult to guarantee equality for all.

VI. Pertubation of public debt

We apply pertubation method to investigate the effects of debt on wealth distribution, young-age income distribution, social welfare and other endogenous variables. This is following Bhandari et al. (2021) in which they decomposed the welfare into three components using pertubation method and discuss how a policy change can affect aggregate efficiency, redistribution and insurance separately. We will give a technical appendix for solving nonlinear equations using comparative static analysis methods (*See appendix C for detail*). We also can use the *backward* command in the pytoch package to quickly calculate the derivative results of individual variables, $\frac{\partial \cdot_1}{\partial \cdot_2}$ denote the coefficient of the derivative of $\cdot_1$ with respect to $\cdot_2$. The specific expression can be found in the section of the corresponding disturbance variables [2].

We only check the effect of the disturbance of debt B in our benchmark model. The superscript, "ˆ", indicates the perturbation term. The results of the comparative static analysis in Appendix C show that the channels of the variables can ultimately be divided into three parts, the direct channel of the constant term of public debt disturbance ($\hat{B}$), and the indirect channel of consumption ratio ($\hat{\rho}$) and investment ratio disturbance ($\hat{\phi}$).

Step 1: Using differential approximation to calculate the perturbation of variables such as $\hat{R}$, $\hat{K}$, $\hat{\tau_w}$, $\hat{\phi}$, $\hat{\rho}$ and so on. We analyze the influence of disturbance on endogenous variables and obtained the perturbation value of the endogenous variable using the differential approximation method, see Table 6.

Step 2: Channel decomposition of $\hat{\tau_w}$, $\hat{R}$, $\hat{K}$, $\omega(\hat{\theta})$, $J(\hat{\theta})$, $\hat{w}$ and $\hat{\bar{w}}$ We take the example of a positive perturbation in the public debt, i.e., an increase in public debt of 0.01. The expressions for the channel decomposition are obtained using comparative static analysis (see Equation (16)-(20)), and the numerical solution of the channel decomposition is derived automatically using pytorch (Table 7).

Overall, the channel decomposition results for the endogenous variables consist of three main aspects: the direct utility channel influenced by the constant term ($\hat{B}$), the indirect utility channel influenced by the consumption ratio ($\hat{\rho}$) and the

[2]See the Appendix C: Equation (C2) for the derivative coefficients for R, $\frac{\partial R}{\partial \cdot}$; Equation (C3) for the derivative coefficients for K, $\frac{\partial K}{\partial \cdot}$; Equation (C7) for the derivative coefficients for $\omega(\theta)$, $\frac{\partial \omega(\theta)}{\partial \cdot}$; Equation C8for the derivative coefficients for w, $\frac{\partial w}{\partial \cdot}$; See the Appendix B.B1 for the derivative coefficients for $w(x)$, $\frac{\partial w(x)}{\partial \cdot}$; See the appendix B.B2 for the derivative coefficients for $g(y)$, $\frac{\partial g(y)}{\partial \cdot}$; See the Appendix B.B3 for the derivative coefficients for V, $\frac{\partial V}{\partial \cdot}$.

TABLE 6—ENDOGENOUS VARIABLE PERTURBATION RESULTS

	T	T-1	$\hat{\cdot}_{t-1}$	T+1	$\hat{\cdot}_{t+1}$
B	1.4164	1.4064	-1	1.4264	1
K	0.0732	0.0725	-0.0671	0.0739	0.0675
ρ	0.2621	0.2627	0.0610	0.2615	-0.0605
R	1.6647	1.6753	1.0571	1.6542	-1.0455
ϕ	0.0491	0.0490	-0.0100	0.0492	0.0100
τ_w	0.6210	0.6334	1.2332	0.6087	-1.2302
w	0.8276	0.8253	-0.2282	0.8299	0.2282
$\overline{w}$	0.3136	0.3026	-1.1043	0.3247	1.1074

Note: We obtained the perturbation values of the endogenous variables using differential approximation, the initial state variable is the result at moment T, where $T+1$ denotes a change of 0.01 units in the positive direction of public debt B, $\Delta = +0.01$ and the perturbation of the variables is calculated as $hat\cdot_{t+1} = [(T+1) - T]/\Delta$; $T-1$ denotes a change of 0.01 units in the opposite direction of public debt B, $\Delta = -0.01$ and the perturbation of the variables is calculated as $hat\cdot_{t-1} = [(T-1) - T]/\Delta$. In the subsequent perturbations, $T-1$ and $T+1$ are defined identically.

TABLE 7—CHANNEL DECOMPOSITION BASED ON THE EFFECTS (T+1)

Channel	$\hat{\tau_w}$	$\hat{R}$	$\hat{K}$	$\hat{w}$	$\hat{\overline{w}}$
Direct utility channels of debt ($\hat{B}$)	-2.21	-1.75	0.05	0.18	1.89
Indirect utility channels of debt ($\hat{\phi}$)	0.20	0.17	0.02	0.05	-0.15
Indirect utility channels of debt ($\hat{\rho}$)	0.77	0.52	0.00	0.00	-0.64
Sum	-1.24	-1.05	0.07	0.23	1.11
Match	-1.23	-1.05	0.07	0.23	1.11

Note: Sum: sum of channels; *Match*: actual perturbation value.

investment ratio ($\hat{\phi}$). Both Table and Figure show the perturbation results and corresponding channel decomposition for an increase of 0.01 unit. There is a good coincidence between the total value of channel and the real value of disturbance, which also indicates the correctness of our comparative static analysis equation.

$$\hat{\tau_w} = \hat{B}\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}} + \hat{\phi}\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\} - \hat{\rho}\frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}. \tag{16}$$

$$\begin{aligned}\hat{R} = {} & \hat{B}\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}}\frac{\partial R}{\partial \tau_w} + \hat{\phi}\left[\frac{\partial R}{\partial \phi} + \frac{\partial R}{\partial \tau_w}\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\}\right] \\ & + \hat{\rho}\left[\frac{\partial R}{\partial \rho} - \frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\frac{\partial R}{\partial \tau_w}\right].\end{aligned} \tag{17}$$

$\hat{\tau_w}$ **&** $\hat{R}$**:** The direct utility channel shows that as the debt increases, the government will have more funds to promote socioeconomic development and stimulate domestic demand, such as increasing infrastructure and employment, both the tax rate and risk-free interest rate will be reduced; the indirect effect channel is somewhat weakened by the reduction in the tax rate.

$$\hat{K} = \hat{B}\frac{\phi}{(1-\phi)} + \hat{\phi}\frac{K}{(1-\phi)\phi}. \tag{18}$$

$$\hat{w} = \hat{B}\frac{\partial w}{\partial K}\frac{\phi}{(1-\phi)} + \hat{\phi}\frac{\partial w}{\partial K}\frac{K}{(1-\phi)\phi}. \tag{19}$$

$$\begin{aligned}\hat{\overline{w}} = \hat{B}&\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{\phi}{(1-\phi)} - w\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}}\right] + \hat{\rho}\left[w\frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\right] \\ &+ \hat{\phi}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{K}{(1-\phi)\phi} - w\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\}\right].\end{aligned} \tag{20}$$

$\hat{K}$ **&** $\hat{w}$ **&** $\hat{\overline{w}}$**:** The direct utility of the channel ($\hat{B}$) shows that the increase of public debt is beneficial to the increase of capital and the wage rate, thus contributing to the increase in after-tax wage rate; And the indirect effect channel ($\hat{\phi}$) also contribute to capital growth. According to Equation (A12), the expression of wage rate is affected only by the endogenous variable K, so the results of channel decomposition of wage rate are the same as capital.

Further, we graphically illustrate the decomposition of the risk interest rate $\hat{\omega(\theta)}$ and the return on capital $\hat{J(\theta)}$.

$$\hat{\omega(\theta)} = \hat{B}\frac{\partial \omega(\theta)}{\partial K}\frac{\phi}{(1-\phi)} + \hat{\phi}\frac{\partial \omega(\theta)}{\partial K}\frac{K}{(1-\phi)\phi}. \tag{21}$$

$\hat{\omega(\theta)}$**:** The perturbation channel of risk interest rate is divided into two parts, this is because expressions of risk interest rates receive only the effect of capital K.

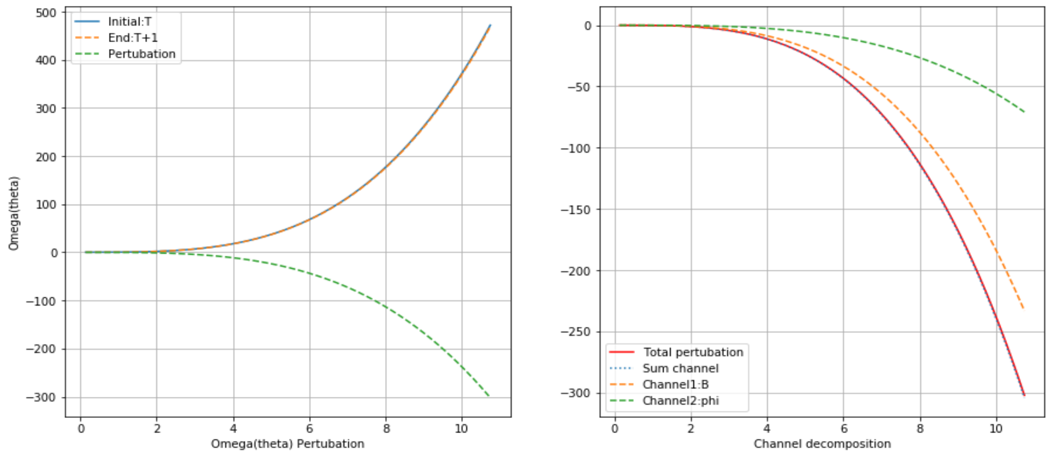

FIGURE 6. CHANNEL DECOMPOSITION OF $\omega(\theta)$[T+1]

Note: Initial:T denotes the initial state; *End:T+1* indicates the end state; *Pertubation* denotes the approximate result of the difference. *Sum channel* represents the sum of all channels. *Total pertubation* represents the true disturbance value.

When public debt changes in a positive direction, leading to an increase in capital inflows into the market, this increase in the money supply is bound to lower its market interest rate, which is why the risk rate disturbance is negative.

(22)

$$\begin{aligned}
J(\hat{\theta}) = \hat{B}&\left[\frac{\partial\omega(\theta)}{\partial K}\frac{\phi^2}{(1-\phi)} + \frac{\phi}{\frac{\partial K}{\partial\tau_w}}\frac{\partial R}{\partial\tau_w}\right] \\
&+ \hat{\phi}\left[\frac{\partial\omega(\theta)}{\partial K}\frac{K}{(1-\phi)} + \omega(\theta) - R + \frac{\partial R}{\partial\phi}(1-\phi) + \frac{\partial R}{\partial\tau_w}\frac{1}{\frac{\partial K}{\partial\tau_w}}\left\{\frac{K}{\phi} - \frac{\partial K}{\partial\phi}(1-\phi)\right\}\right] \\
&+ \hat{\rho}\left[\frac{\partial R}{\partial\rho} - \frac{1}{\frac{\partial K}{\partial\tau_w}}\frac{\partial K}{\partial\rho}\frac{\partial R}{\partial\tau_w}\right](1-\phi).
\end{aligned}$$

$J(\hat{\theta})$**:** Among them, direct effect channels have the greatest impact when increased public debt leads to a decline in capital yields due to the corresponding reduction in risk interest rates. In addition, the increase in public debt is conducive to stimulating domestic demand, and the positive effects of indirect utility channels partially weaken the negative effects of direct effects.

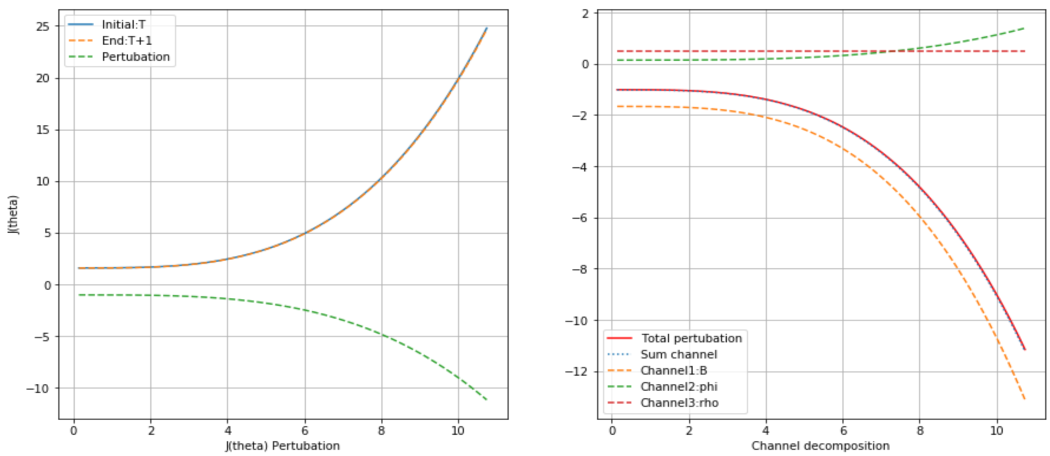

FIGURE 7. CHANNEL DECOMPOSITION OF $J(\theta)$[T+1]

Note: Initial:T denotes the initial state; *End:T+1* indicates the end state; *Pertubation* denotes the approximate result of the difference. *Sum channel* represents the sum of all channels. *Total pertubation* represents the true disturbance value.

Step 3: Channel decomposition of stable distribution Unlike previous literature, we have found a stable distribution of wealth and young-age income using the principle of machine learning, the corresponding technical steps to stabilize the distribution map and loss function are set out in the Appendix B.B4. Then we use interpolation method to analyze the disturbance of different distributions.

$$\begin{aligned}
\hat{w(x)} &= \hat{w}\left(\nu\right)\frac{\partial w(x)}{\partial w\left(\nu\right)} + \hat{q}\left(\eta\right)\frac{\partial w(x)}{\partial q\left(\eta\right)} \\
&+ \hat{B}\frac{\partial w(x)}{\partial \overline{w}}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{\phi}{(1-\phi)} - w\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}}\right] \\
&+ \hat{\phi}\frac{\partial w(x)}{\partial \overline{w}}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{K}{(1-\phi)\phi} - w\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\}\right] \\
&+ \hat{\rho}\left[\frac{\partial w(x)}{\partial \overline{w}}w\frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho} + \frac{\partial w(x)}{\partial \rho}\right].
\end{aligned} \tag{23}$$

$\hat{w(x)}$**:** Where $q(\cdot)$ is the probability density function for labor shocks η, $p(\theta)$ is the probability density function of θ, $w(\cdot)$ is the probability density function of wealth.

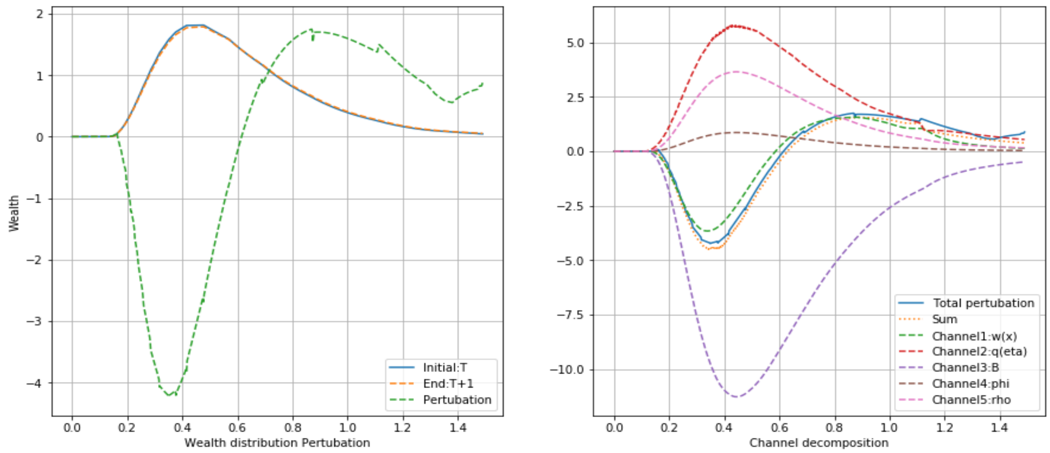

FIGURE 8. CHANNEL DECOMPOSITION OF WEALTH DISTRIBUTION $w(x)$ [T+1]

Among them, we directly express some variable coefficients by the definition of their derivatives, for example, $\dfrac{\partial w(x)}{\partial \overline{w}}$ refers to the coefficient obtained after taking the derivative of the distribution $w(x)$ with respect to $\overline{w}$.

The stable distribution of wealth has five channels. The first term is the indirect channel of x which describes the effect of changes in debt on wealth itself; The second term is the pertubation of the statistical density function of η, which describes the effect of the change of debt on the labor shock. Other channels include constant items of change in public debt $\hat{B}$, investment channels $\hat{\phi}$ and consumption channels $\hat{\rho}$ (See Figure 8). Of the five channels, public debt channels have the greatest impact and investment channels have the smallest impact. The increase in the public debt will directly lead to a reduction in wealth, because the issuance of public debt is the state borrowing money from the people where, which is equivalent to reducing the value of wealth in the hands of the people. But then, after the issuance of the public debt, the country has money to build infrastructure, ah, or some transfer payments, as well as the generation of some investment projects, will boost the country's domestic demand, that is, the positive impact of the investment and consumption channel ratio here, the final impact, although still negative, but the investment channel these to some extent or weaken this part of the results.

Proof. See Appendix B.B1

$$
\begin{aligned}
\hat{g(y)} &= g\,(\hat{\mu})\frac{\partial g(y)}{\partial g\,(\mu)} + \hat{q}\,(\eta)\,\frac{\partial g(y)}{\partial q\,(\eta)} \\
&+ \hat{B}\frac{\partial g(y)}{\partial \overline{w}}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{\phi}{(1-\phi)} - w\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}}\right] \\
&+ \hat{\phi}\frac{\partial g(y)}{\partial \overline{w}}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{K}{(1-\phi)\phi} - w\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\}\right] \\
&+ \hat{\rho}\frac{\partial g(y)}{\partial \overline{w}}\left[w\frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\right].
\end{aligned} \tag{24}
$$

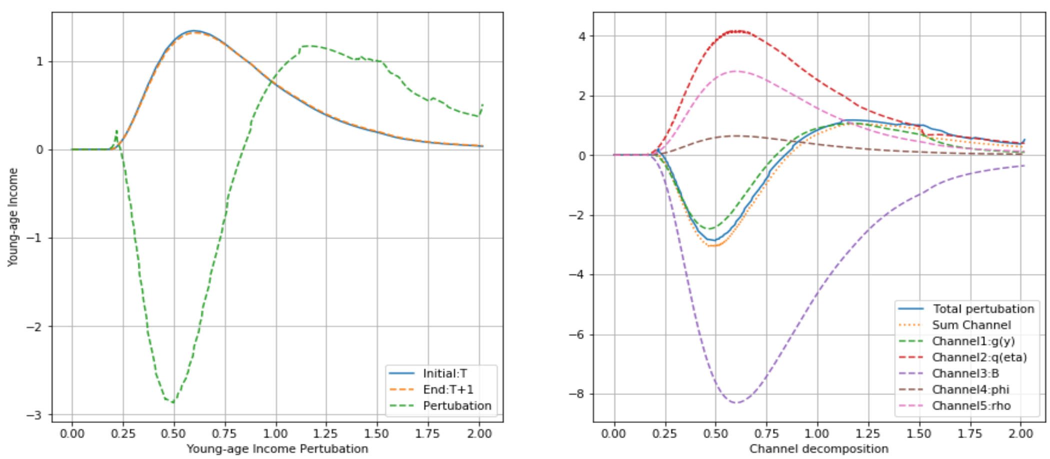

FIGURE 9. CHANNEL DECOMPOSITION OF YOUNG-AGE INCOME DISTRIBUTION $g(y)$ [T+1]

$\hat{g(y)}$**:** Where $q(\cdot)$ is the probability density function for labor shocks η, $p(\theta)$ is the probability density function of θ, $g(\cdot)$ is the probability density function of young-age income. Since young-age income y and wealth x differ only by a factor of $(1-\rho)$, the channel decomposition of young-age income is essentially the same as wealth (See Figure 9).

Proof. See Appendix B.B2

Step 4: Channel decomposition of welfare $\hat{V}$ Based on the previous analysis, we fully characterize the equilibrium for a given exemption level, it is now possible to understand how changes in the optimal level B affect social welfare. Because firm make zero profit in equilibrium, individuals maximizing indirect utility $V(B)$, maximizes social welfare in this economy.

Lemma VI.1 (Directional test for a change in the public debt level** B**) *The effect of a debt reform $\hat{B}$ of the initial debt B on social welfare, $\hat{V}$, is the solution to the functional equation: for all $\theta \in \Theta$, the normalized welfare change induced by a marginal change in the public debt B is given by:*

(25)

$$\begin{aligned}\hat{V} = {} & \hat{B}\frac{\partial V}{\partial J(\theta)}E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\frac{\partial\omega(\theta)}{\partial K}\frac{\phi}{(1-\phi)}\phi + \frac{\phi}{\frac{\partial K}{\partial \tau_w}}\frac{\partial R}{\partial \tau_w}\right]\right\} \\ & + \hat{\phi}\frac{\partial V}{\partial EJ(\theta)}E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\begin{array}{c}\frac{\partial\omega(\theta)}{\partial K}\frac{K}{(1-\phi)} + \omega(\theta) - R + \frac{\partial R}{\partial \phi}(1-\phi) \\ +\frac{\partial R}{\partial \tau_w}\frac{1}{\frac{\partial K}{\partial \tau_w}}\left(\frac{K}{\phi} - \frac{\partial K}{\partial \phi}(1-\phi)\right)\end{array}\right]\right\} \\ & + \hat{\rho}\frac{\partial V}{\partial EJ(\theta)}E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\frac{\partial R}{\partial \rho} - \frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\frac{\partial R}{\partial \tau_w}\right](1-\phi)\right\} \\ & + g\hat{(y)}\frac{\partial V}{\partial g(y)},\end{aligned}$$

$\hat{V}$**:** Lemma VI.1 , which presents the central result of this paper, provides a test for whether to optimally increase or decrease the public debt and implies that measures of four observable or recoverable variables are sufficient to determine whether the level of public debt is optimal or should be increased or decreased. In the Equation (25), we show that the marginal welfare change caused by the change in public debt level B corresponds to the following expression, where $g\hat{(y)}$ represents the effect of debt change on young-age income distribution; $\hat{\phi}$ represents the effect of debt change on the investment ratio; $\hat{\rho}$ represents the effect of debt change on the consumption-income ratio; and $\hat{B}$ is either 1 or -1. Direct effect channels ($\hat{B}$) have the greatest impact. Indirect effect channels ($\hat{\phi}$) have the smallest impact.

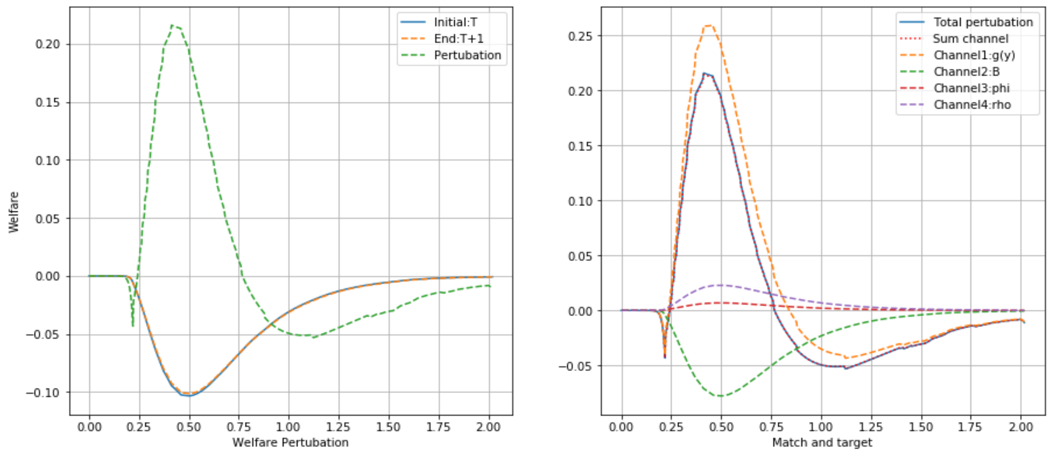

FIGURE 10. CHANNEL DECOMPOSITION OF WELFARE DISTRIBUTION V [T+1]

Proposition VI.1 (Sufficient statistics for optimal public debt B***)*** *Through the characterization of welfare changes in VI.1, the optimal public debt level theory is described as follows Dávila (2020):*

(26)

$$B^* = \frac{E\left\{J(\theta)^{-\gamma}\left[\begin{array}{c} -\hat{B}\underbrace{\left[K(\omega(\theta)-1+\delta)\right]}_{>0} \\ -\epsilon_{\phi,B}K^2\underbrace{\left[\frac{\partial\omega(\theta)}{\partial K}\frac{1}{(1-\phi)}+\frac{1}{\phi}\left(\omega(\theta)-R+\frac{\frac{\partial R}{\partial\tau_w}}{\frac{\partial K}{\partial\tau_w}}\right)\right]}_{<0}\end{array}\right]\right\}}{E\left\{J(\theta)^{-\gamma}\left[\begin{array}{c}\hat{B}\,\underbrace{K\frac{\frac{\partial R}{\partial\tau_w}}{\frac{\partial K}{\partial\tau_w}}}_{<0}+\epsilon_{\phi,B}\phi\underbrace{\left[\frac{\partial R}{\partial\phi}-\frac{\frac{\partial R}{\partial\tau_w}}{\frac{\partial K}{\partial\tau_w}}\frac{\partial K}{\partial\phi}\right]}_{>0} \\ +\epsilon_{\rho,B}\rho\underbrace{\left[\frac{\partial R}{\partial\rho}-\frac{\frac{\partial R}{\partial\tau_w}}{\frac{\partial K}{\partial\tau_w}}\frac{\partial K}{\partial\rho}\right]}_{<0}\end{array}\right]\right\}+\frac{\partial g(y)}{\partial B}\underbrace{\frac{\frac{\partial V}{\partial g(y)}}{(1-\gamma)\frac{\partial V}{\partial EJ(\theta)}}\frac{K}{\phi}}_{<0}}$$

Where, $\hat{\cdot}$ *represents the disturbance term of each channel.* $\frac{\partial\cdot_1}{\partial\cdot_2}$ *denote the coefficient of the derivative of* $\cdot_1$ *with respect to* $\cdot_2$*,* ϵ *denotes the coefficient of elasticity.*

Proof. *See Appendix B.B3*

In the Equation (26), if the public debt increases by 1 unit, the disturbance value, $\hat{B} = 1$, and the distributive disturbance value of young income $g(\hat{y}) < 0$. At this time, the public debt channel in the numerator is less than 0, and the rest is greater than 0; The denominator is greater than 0 for all but the public debt channel.

If the public debt decreases by 1 unit, the disturbance value, $\hat{B} = -1$, and the distributive disturbance value of young income $g(\hat{y}) > 0$. In this case, all the numerators are greater than 0, and all the denominators are greater than 0 except for the young-age income channel.

In general, direct utility channels are oriented opposite to indirect utility channels. The determination of the optimal level of national debt is a game between the direct channel and the indirect channel, where the value of welfare decreases when the direct channel is more favorable and increases when the indirect channel is more favorable. Eventually the optimal level of national debt is obtained when the two are equivalent.

VII. Conclusion

Motivated by the dramatic surge in public debt/GDP ratios in the COVID-19 health crisis period, our focus is mainly on the relationship between debt and inequality. In this paper, by targeting the quantities of wealth and earnings distribution of the US economy in our calibration, we use macro modeling, a two-period OLG model with idiosyncratic investment risk, proposed a novel technique to identify and characterize several recent secular trends: the increase in income inequality, the optimal public debt and welfare, the decline in natural interest rates, and the rise in debt by households and governments.

The issue of public debt affects the return on assets in the market, and then affects the macroeconomic equilibrium and wealth distribution. Public debt also provide funding liquidity in financial markets.

From the empirical point of view, we have calculated the optimum quantity of debt for a model that is parameterized to mimic certain features of the US economy.

From a normative point of view, on the one hand, we use welfare maximization to argue what is optimal debt; We find that there is an inverted U-shape of welfare that first rises and then falls, with the corresponding Gini coefficient showing a U-shape. Moreover, when welfare is the largest, social inequality is the smallest. On the other hand, through the comparative static analysis of the stable distribution, the influence results of the disturbance problem under the optimal debt level are analyzed. We use the first-order condition on welfare to obtain a sufficient statistical form for the optimal debt. Channel decomposition is carried out for wealth distribution, welfare changes and corresponding endogenous variables. In general, a series of changes brought about by changes in public debt are ultimately caused by the corresponding changes in the investment ratio and the consumption ratio.

There are three contributions to the article. (1) We use a two-period overlapping generations model to investigate the relationship between government bonds and wealth inequality. (2) We use machine learning to investigate the impacts of debts on the economy through different channels. (3) We find the optimal debt level in terms of the social welfare maximization.

REFERENCES

Acikgoz, Omer. 2015. "Transitional dynamics and long-run optimal taxation under incomplete markets."

Acikgoz, Omer, Marcus Hagedorn, Hans Aasnes Holter, and Yikai Wang. 2018. "The optimum quantity of capital and debt." *Available at SSRN 3185088.*

Aiyagari, S Rao. 1994. "Uninsured idiosyncratic risk and aggregate saving." *The Quarterly Journal of Economics*, 109(3): 659–684.

Aiyagari, S Rao. 1995. "Optimal capital income taxation with incomplete markets, borrowing constraints, and constant discounting." *Journal of political Economy*, 103(6): 1158–1175.

Aiyagari, S Rao, and Ellen R McGrattan. 1998. "The optimum quantity of debt." *Journal of Monetary Economics*, 42(3): 447–469.

Alonso-Ortiz, Jorge, and Richard Rogerson. 2010. "Taxes, transfers and employment in an incomplete markets model." *Journal of Monetary Economics*, 57(8): 949–958.

Attanasio, Orazio P, James Banks, Costas Meghir, and Guglielmo Weber. 1999. "Humps and bumps in lifetime consumption." *Journal of Business & Economic Statistics*, 17(1): 22–35.

Barro, Robert J. 1979. "On the determination of the public debt." *Journal of political Economy*, 87(5, Part 1): 940–971.

Bewley, Truman. 1986. "Stationary monetary equilibrium with a continuum of independently fluctuating consumers." *Contributions to mathematical economics in honor of Gérard Debreu*, 79.

Bhandari, Anmol, David Evans, Mikhail Golosov, and Thomas Sargent. 2021. "Efficiency, Insurance, and Redistribution Effects of Government Policies." working paper.

Brock, William A, and Leonard J Mirman. 1972. "Optimal economic growth and uncertainty: the discounted case." *Journal of Economic Theory*, 4(3): 479–513.

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull. 2003. "Accounting for the US earnings and wealth inequality." *Journal of political economy*, 111(4): 818–857.

Chang, Roberto. 2022. "Should Central Banks Have an Inequality Objective?" National Bureau of Economic Research.

Dávila, Eduardo. 2020. "Using elasticities to derive optimal bankruptcy exemptions." *The Review of Economic Studies*, 87(2): 870–913.

De Nardi, Mariacristina. 2004. "Wealth inequality and intergenerational links." *The Review of Economic Studies*, 71(3): 743–768.

Díaz-Giménez, Javier, Andy Glover, and José-Víctor Ríos-Rull. 2011. "Facts on the distributions of earnings, income, and wealth in the United States: 2007 update." *Federal Reserve Bank of Minneapolis Quarterly Review*, 34(1): 2–31.

Dyrda, Sebastian, and Marcelo Pedroni. 2021. "Optimal Fiscal Policy in a Model with Uninsurable Idiosyncratic Income Risk." *Available at SSRN 3289306*.

Flodén, Martin. 2001. "The effectiveness of government debt and transfers as insurance." *Journal of Monetary Economics*, 48(1): 81–108.

Gottardi, Piero, Atsushi Kajii, and Tomoyuki Nakajima. 2015. "Optimal taxation and debt with uninsurable risks to human capital accumulation." *American Economic Review*, 105(11): 3443–70.

Gourinchas, Pierre-Olivier, and Jonathan A Parker. 2002. "Consumption over the life cycle." *Econometrica*, 70(1): 47–89.

Kindermann, Fabian, and Dirk Krueger. 2014. “High marginal tax rates on the top 1%? Lessons from a life cycle model with idiosyncratic income risk.” National Bureau of Economic Research.

Mian, Atif, Ludwig Straub, and Amir Sufi. 2021. “Indebted demand.” *The Quarterly Journal of Economics*, 136(4): 2243–2307.

Nakajima, Tomoyuki, and Shuhei Takahashi. 2017. “The optimum quantity of debt for Japan.” *Journal of the Japanese and International Economies*, 46: 17–26.

Prescott, Edward C. 1986. “Theory ahead of business-cycle measurement.” Vol. 25, 11–44, Elsevier.

Röhrs, Sigrid, and Christoph Winter. 2017. “Reducing government debt in the presence of inequality.” *Journal of Economic Dynamics and Control*, 82: 1–20.

Wan, Jing, and Shenghao Zhu. 2019. “Bequests, estate taxes, and wealth distributions.” *Economic Theory*, 67(1): 179–210.

Technical appendix of the benchmark model

A1. Environment

Proof of Personal Utility Maximization–Old-age (Proposition II.1)

1. (l_{t+1}, π_{t+1}) solves $\max\limits_{l_{t+1}} \theta_{t+1} A k_{t+1}^{\alpha} l_{t+1}^{1-\alpha} - w_{t+1} l_{t+1}$

$$\frac{\partial \pi_{t+1}}{\partial l_{t+1}} = \theta_{t+1} A(1-\alpha)\left(\frac{k_{t+1}}{l_{t+1}}\right)^{\alpha} - w_{t+1} = 0,$$

$$w_{t+1} = \theta_{t+1} A(1-\alpha)\left(\frac{k_{t+1}}{l_{t+1}}\right)^{\alpha},$$

$$l_{t+1} = \left[\frac{\theta_{t+1} A(1-\alpha)}{w_{t+1}}\right]^{\frac{1}{\alpha}} k_{t+1}. \tag{A1}$$

$$\pi_{t+1} = \theta_{t+1} A k_{t+1}^{\alpha}\left(\left[\frac{\theta_{t+1} A(1-\alpha)}{w_{t+1}}\right]^{\frac{1}{\alpha}} k_{t+1}\right)^{1-\alpha} - w_{t+1}\left(\frac{\theta_{t+1} A(1-\alpha)}{w_{t+1}}\right)^{\frac{1}{\alpha}} k_{t+1},$$

$$\pi_{t+1} = (\theta_{t+1} A)^{\frac{1}{\alpha}} k_{t+1}\left(\frac{1-\alpha}{w_{t+1}}\right)^{\frac{1-\alpha}{\alpha}} - (\theta_{t+1} A)^{\frac{1}{\alpha}} k_{t+1}\frac{(1-\alpha)^{\frac{1}{\alpha}}}{(w_{t+1})^{\frac{1-\alpha}{\alpha}}},$$

$$\pi_{t+1}(\theta_{t+1}, k_{t+1}) = \alpha(\theta_{t+1} A)^{\frac{1}{\alpha}}\left(\frac{w_{t+1}}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}} k_{t+1}. \tag{A2}$$

2. $(c_{2,t+1},\ z_{t+1})$ solves

$$\max_{c_{2,t+1}; z_{t+1}} \frac{c_{2,t+1}^{1-\gamma}}{1-\gamma} + \chi\frac{[(1-\tau_{z,t+1}) z_{t+1}]^{1-\gamma}}{1-\gamma}, \text{ s.t.} \quad c_{2,t+1} + z_{t+1} = h_{t+1};$$

The first-order condition and the second utility maximization can be obtained,

$$z_{t+1} = h_{t+1}\frac{\chi^{\frac{1}{\gamma}}(1-\tau_{z,t+1})^{\frac{1-\gamma}{\gamma}}}{1+\chi^{\frac{1}{\gamma}}(1-\tau_{z,t+1})^{\frac{1-\gamma}{\gamma}}}, \tag{A3}$$

$$c_{2,t+1} = h_{t+1}\frac{1}{1+\chi^{\frac{1}{\gamma}}(1-\tau_{z,t+1})^{\frac{1-\gamma}{\gamma}}}. \tag{A4}$$

3. (h_{t+1}) solves $h_{t+1} = (1-\tau_{p,t+1})\pi_{t+1}(\theta_{t+1}, k_{t+1}) + (1-\delta)k_{t+1} + R_{t+1} b_{t+1}$

$$h_{t+1} = k_{t+1}\left[(1-\tau_{p,t+1})\,\alpha\,(\theta_{t+1} A)^{\frac{1}{\alpha}}\left(\frac{w_{t+1}}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}} + 1 - \delta\right] + R_{t+1} b_{t+1},$$

substituting the expression for firm profits, Equation (A2), into the expression for old-age income, we can obtain the income in old-age formed by a combination of the risky and the risk-free interest rates, and the expression for the risky rate of interest on capital, $\omega(\theta)$.

$$h_{t+1} = k_{t+1}\omega_{t+1}(\theta_{t+1}) + R_{t+1}b_{t+1}, \tag{A5}$$

$$\omega_{t+1}(\theta_{t+1}) = (1-\tau_{p,t+1})\,\alpha\,(\theta_{t+1}A)^{\frac{1}{\alpha}}\left(\frac{w_{t+1}}{1-\alpha}\right)^{\frac{\alpha-1}{\alpha}} + 1 - \delta, \tag{A6}$$

4. $(u(c_{2,t+1}, z_{t+1}))$ The second question can be written in the following form,

$$\max_{c_{2,t+1};z_{t+1}} \frac{c_{2,t+1}^{1-\gamma}}{1-\gamma} + \chi\frac{[(1-\tau_{z,t+1})z_{t+1}]^{1-\gamma}}{1-\gamma} = \left(1+\chi^{\frac{1}{\gamma}}(1-\tau_{z,t+1})^{\frac{1-\gamma}{\gamma}}\right)^{\gamma}\frac{h_{t+1}^{1-\gamma}}{1-\gamma},$$

$$u(c_{2,t+1}, z_{t+1}) = \frac{1}{1-\gamma}\Gamma\left[\omega_{t+1}(\theta_{t+1})k_{t+1} + R_{t+1}b_{t+1}\right]^{1-\gamma}, \tag{A7}$$

$$\Gamma = \left(1+\chi^{\frac{1}{\gamma}}\left(1-\tau_{z,t+1}\right)^{\frac{1-\gamma}{\gamma}}\right)^{\gamma}. \tag{A8}$$

Proof of Personal Utility Maximization–Young-age (Proposition II.2)

1. $(c_{1,t})$ solves $\max_{c_{1,t}}\frac{c_{1,t}^{1-\gamma}}{1-\gamma} + \beta EU(c_{2,t+1}, z_{t+1})$, s.t. $\quad c_{1,t} + k_{t+1} + b_{t+1} = y_t$;

Denote $\rho_t = c_{1,t}/y_t$ to be the expenditure share on consumption in the young-age. The remaining expenditure then go to investment. Denote $\phi_{t+1} = k_{t+1}/(k_{t+1}+b_{t+1})$ to be the fraction of capital in the investment portfolio.

2. $(\rho_t,\, \phi_{t+1})$ solves

$$\max_{\rho_t;\phi_{t+1}} \frac{(\rho_t y_t)^{1-\gamma}}{1-\gamma} + \beta\Gamma\left[y_t(1-\rho_t)\right]^{1-\gamma} E\left\{\frac{\left[\omega_{t+1}(\theta_{t+1})\phi_{t+1} + R_{t+1}(1-\phi_{t+1})\right]^{1-\gamma}}{1-\gamma}\right\}$$

$$\frac{\partial L(\rho_t, \phi_{t+1})}{\partial \phi_{t+1}} = 0:$$

$$E\left\{\left[\omega_{t+1}(\theta_{t+1})\phi_{t+1} + R_{t+1}(1-\phi_{t+1})\right]^{-\gamma}\left(\omega_{t+1}(\theta_{t+1}) - R_{t+1}\right)\right\} = 0, \tag{A9}$$

$$\frac{\partial L(\rho_t, \phi_{t+1})}{\partial \rho_t} = 0 :$$

$$\text{(A10)} \qquad \rho_t = \frac{\left\{E[\omega_{t+1}(\theta_{t+1})\phi_{t+1} + R_{t+1}(1-\phi_{t+1})]^{1-\gamma}\right\}^{-\frac{1}{\gamma}}}{(\beta\Gamma)^{\frac{1}{\gamma}} + \left\{E[\omega_{t+1}(\theta_{t+1})\phi_{t+1} + R_{t+1}(1-\phi_{t+1})]^{1-\gamma}\right\}^{-\frac{1}{\gamma}}}.$$

3. $(u(y_t))$ solves

$$\frac{y_t^{1-\gamma}}{1-\gamma}\left\{\rho_t^{1-\gamma} + \beta\Gamma(1-\rho_t)^{1-\gamma} E\left\{[\omega_{t+1}(\theta_{t+1})\phi_{t+1} + R_{t+1}(1-\phi_{t+1})]^{1-\gamma}\right\}\right\}$$

$$\text{(A11)} \qquad u(y_t) = \frac{y_t^{1-\gamma}}{1-\gamma}\left\{\rho_t^{1-\gamma} + \beta\Gamma(1-\rho_t)^{1-\gamma} E\left\{J(\theta)^{1-\gamma}\right\}\right\}$$

Therefore, the utility equation and the optimal indirect utility function of the individual can be rewritten as above.

A2. General equilibrium solutions

Proof of the Steady State (Proposition II.3)

1. (w) solves $\int l = 1$

$$\int l_{t+1} = \int \left(\frac{\theta_{t+1}A(1-\alpha)}{w_{t+1}}\right)^{\frac{1}{\alpha}} k_{t+1} = 1, \int [\theta_{t+1}A(1-\alpha)]^{\frac{1}{\alpha}} k_{t+1} = w_{t+1}^{\frac{1}{\alpha}}$$

$$\text{(A12)} \qquad w = (1-\alpha)AK^{\alpha}\left(E\theta^{\frac{1}{\alpha}}\right)^{\alpha}.$$

2. (k_t) solves $k_{t+1} = \phi_{t+1}(k_{t+1} + b_{t+1})$ and $\varphi = \dfrac{\chi^{\frac{1}{\gamma}}(1-\tau_z)^{\frac{1}{\gamma}}}{1+\chi^{\frac{1}{\gamma}}(1-\tau_z)^{\frac{1-\gamma}{\gamma}}}$

$$\text{(A13)} \qquad k_{t+1} = \phi_{t+1}(1-\rho_t)\left\{\begin{array}{c}\left[\begin{array}{c}\alpha A(1-\tau_{pt})\theta_t^{\frac{1}{\alpha}}\left(E\theta_t^{\frac{1}{\alpha}}\right)^{\alpha-1} K_t^{\alpha-1} \\ +1-\delta+R_t\left(\dfrac{1-\phi_t}{\phi_t}\right)\end{array}\right] \\ *\varphi k_t + (1-\tau_{wt})\eta_t(1-\alpha)AK_t^{\alpha}\left(E\theta_t^{\frac{1}{\alpha}}\right)^{\alpha}\end{array}\right\}.$$

In the derivation of the above equation, we use the definition of the investment ratio ϕ, the individual's budget constraint, the consumption ratio ρ, the after-tax inheritance income ratio φ; the expression for the income h_t at old-age (Equation A5), the bequest z_t (Equation A3), the wage rate w (Equation A12), and the risky interest rate $\omega_{t+1}(\theta)$ (Equation A6).

3. (K_t) solves $\int k_{t+1} = K_{t+1}$ with Equation (A13)

$$K_{t+1} = \phi_{t+1}(1-\rho_t)\left\{\begin{array}{c}\varphi K_t\left[\begin{array}{c}\alpha A(1-\tau_{pt})\left(E\theta_t^{\frac{1}{\alpha}}\right)^{\alpha}K_t^{\alpha-1}\\ +1-\delta+R_t\left(\dfrac{1-\phi_t}{\phi_t}\right)\end{array}\right]\\ +(1-\tau_{wt})(1-\alpha)AK_t^{\alpha}(E\theta_t^{\frac{1}{\alpha}})^{\alpha}\end{array}\right\} \tag{A14}$$

4. (K) solves $K_{t+1} = K_t = K, \phi_t = \phi_{t+1} = \phi$

$$K^{\alpha-1} = \frac{\dfrac{1}{\phi(1-\rho)} - \varphi(1-\delta) - R\varphi\left(\dfrac{1-\phi}{\phi}\right)}{A\left(E\theta_t^{\frac{1}{\alpha}}\right)^{\alpha}\{\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)\}}. \tag{A15}$$

The formula states that capital is jointly determined by the risk-free interest rate R, the consumption ratio ρ, the investment ratio ϕ and the tax rate τ_w.

5. (H_t) solves $H_t = \int h_t = \int \omega_t(\theta_t)k_t + R_t b_t$

$$H_t = K\left[\frac{\alpha(1-\tau_p)\left[\dfrac{1}{\phi(1-\rho)} - \varphi(1-\delta) - R\varphi\left(\dfrac{1-\phi}{\phi}\right)\right]}{\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)} + 1 - \delta + \left(\frac{1-\phi}{\phi}\right)R\right]$$

The derivation of the above equation applies the following definitions: the definition of the risky rate of interest $\omega_{t+1}(\theta_{t+1})$ (Equation A6), and capital K in steady state (Equation A15).

6. (R_t) solves $R_tB_t = B_{t+1} + \tau_{wt}w_t + \tau_{zt}\int z_t + \tau_{pt}\int \pi_{t+1}(\theta_{t+1}, k_{t+1})$

$$R_t(1-\phi_t)X_t = (1-\phi_{t+1})X_{t+1} + \tau_{wt}w_t + \frac{\varphi\tau_{zt}H_t}{1-\tau_z} + \tau_{pt}\alpha A\left(E\theta_{t+1}^{\frac{1}{\alpha}}\right)^{\alpha} K_{t+1}^{\alpha} \tag{A16}$$

Define $x_{t+1} = k_{t+1} + b_{t+1}$ as individual's wealth, we can get the relationship between public debt and wealth, $b_{t+1} = (1-\phi_{t+1})x_{t+1}$. The derivation of the above equation applies the following definitions: the definition of the public debt market clear; The government decides; The definition of total profit, heritage, z_t (Equation A3), individual's wealth x_t and the definition of the aggregate income.

7. (R) solves $R_t = R_{t+1}$

$$R\frac{1-\phi}{\phi}K = \frac{1-\phi}{\phi}K + \tau_w w + \tau_z H\frac{\varphi}{1-\tau_z} + \tau_p\alpha A\left(E\theta_{t+1}^{\frac{1}{\alpha}}\right)^{\alpha} K^{\alpha},$$

(A17)

$$R = \frac{1 + \dfrac{\left[\begin{array}{l}\dfrac{1}{(1-\phi)(1-\rho)}\left[(1-\alpha)\tau_w + \tau_p\alpha + \dfrac{\tau_z\varphi}{1-\tau_z}\alpha(1-\tau_p)\right] \\ + \dfrac{(1-\delta)\varphi\phi}{1-\phi}\left(\dfrac{\tau_z}{1-\tau_z}(1-\tau_w)(1-\alpha) - (1-\alpha)\tau_w - \tau_p\alpha\right)\end{array}\right]}{\varphi\alpha(1-\tau_p) + (1-\tau_w)(1-\alpha)}}{1 + \dfrac{\varphi}{\varphi\alpha(1-\tau_p) + (1-\tau_w)(1-\alpha)}\left[(1-\alpha)\tau_w + \tau_p\alpha - \frac{\tau_z}{1-\tau_z}(1-\alpha)(1-\tau_w)\right]}.$$

The derivation of the above formula uses the following definitions: the expression of wealth x; The expression for the wage rate w; The expression of total income H_t; The expression of capital K in steady state (Equation A15).

8. **Algorithm see Table D1**

Technical appendix for fixed point probability density distribution

B1. Wealth distribution

Proof of the Law of Motion for Wealth

1. (x_{t+1}) solves $x_{t+1} = k_{t+1} + b_{t+1}$ and $k_{t+1} + b_{t+1} + c_t = y_t$

$$\begin{aligned} x_{t+1} &= k_{t+1} + b_{t+1} = y_t - c_t = y_t(1-\rho_t) \\ &= [(1-\tau_{zt})z_t + (1-\tau_{wt})w_t\eta_t](1-\rho_t) \end{aligned}$$

$$= \left[(1-\tau_{zt})h_t \frac{\chi^{\frac{1}{\gamma}}(1-\tau_{zt})^{\frac{1-\gamma}{\gamma}}}{1+\chi^{\frac{1}{\gamma}}(1-\tau_{zt})^{\frac{1-\gamma}{\gamma}}} + (1-\tau_{wt})w_t\eta_t\right](1-\rho_t) = ...$$

And, the dynamics of wealth can be summarized as follows:

$$x_{t+1} = x_t(1-\rho)J(\theta)\varphi + (1-\rho)\overline{w}\eta. \tag{B1}$$

The derivation of the above equation uses the following definitions or formulas: the definition of wealth x_t; The individual's budget constraint expression; The expression of consumption ratio ρ; The expression for bequest z_t (Equation A3); And the after-tax inheritance income ratio φ. The expression of return on capital $J(\theta)$.

Proof of the Stationary Distribution as a Fixed Point (Definition III.1)

1. $(W_{X_t}(x))$ solves $W_{(x,t+1)}(x) = P\{x_{t+1} \leq x\}$, the cumulative distribution function of x_{t+1}, $W_{(x,t+1)}(x)$, is:

$W_{(x,t+1)}(x) = P\{x_{t+1} \leq x\} = P\{x_t(1-\rho)J(\theta)\varphi + (1-\rho)\overline{w}\eta \leq x\}$,

$W_{(x,t+1)}(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_{(y,t)}\left(\frac{x-(1-\rho)\overline{w}\eta}{(1-\rho)\varphi J(\theta)}\right)p(\theta)q(\eta)d\theta d\eta$,

Let, $\nu = \frac{x-(1-\rho)\overline{w}\eta}{(1-\rho)\varphi J(\theta)}$, so $\eta = \frac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}$,

$$W_{t+1}(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} W_{(x,t)}\left(\frac{x-(1-\rho)\overline{w}\eta}{(1-\rho)\varphi J(\theta)}\right)p(\theta)q(\eta)d\theta d\eta \tag{B2}$$

2. $(w_{X_t}(x))$ solves $\frac{\partial W_{X_t}(x)}{\partial x}$, for all x ≥ 0, the probability density function of x_{t+1} is:

$$w_{X_{t+1}}(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w_{X_t}\left(\frac{x-(1-\rho)\overline{w}\eta}{(1-\rho)\varphi J(\theta)}\right)\frac{1}{(1-\rho)\varphi J(\theta)}p(\theta)q(\eta)\,d\theta d\eta, \tag{B3}$$

3. $(w(x))$ solves $w_{X_t}(x) = w_{X_{t+1}}(x)$ for all $x \geq 0$, the stationary distribution of $\{x_t\}_{t=0}^{\infty}$ satisfies

$$w_X(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w_X(\nu)\,p(\theta)q\left(\frac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\frac{1}{(1-\rho)\overline{w}}d\theta d\nu. \tag{B4}$$

Proof of the Pertubation on Wealth Distribution The stable distribution of wealth obeys the probability density function $w(x)$ and cumulative distribution function $W(x)$. The effect of a debt change $\hat{B}$ of initial debt on the distribution of individual wealth, $\hat{w(x)}$, is the solution to the equation.

1. $(\hat{w(x)})$ solves $\dfrac{\partial w(x)}{\partial B}$ with Equation (C9), for all $\theta \in \Theta$

$$
\begin{aligned}
\hat{w(x)} = &\textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{w}\,(\nu)\,p(\theta)q\left(\dfrac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\dfrac{1}{(1-\rho)\overline{w}}d\theta d\nu \\
&+\textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\,(\nu)\,p(\theta)\hat{q}\left(\dfrac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\dfrac{1}{(1-\rho)\overline{w}}d\theta d\nu \\
&+\textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\,(\nu)\,p(\theta)q\left(\dfrac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\dfrac{1}{(1-\rho)^2\overline{w}}\hat{\rho}d\theta d\mu \\
&-\textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\,(\nu)\,p(\theta)q\left(\dfrac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\dfrac{1}{(1-\rho)(\overline{w})^2}\hat{B}\dfrac{\partial\hat{\overline{w}}}{\partial\hat{B}}d\theta d\mu \\
&-\textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\,(\nu)\,p(\theta)q\left(\dfrac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\dfrac{1}{(1-\rho)(\overline{w})^2}\hat{\phi}\dfrac{\partial\hat{\overline{w}}}{\partial\hat{\phi}}d\theta d\mu \\
&-\textstyle\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w\,(\nu)\,p(\theta)q\left(\dfrac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\dfrac{1}{(1-\rho)(\overline{w})^2}\hat{\rho}\dfrac{\partial\hat{\overline{w}}}{\partial\hat{\rho}}d\theta d\mu,
\end{aligned}
$$

$$
\begin{aligned}
\hat{w(x)} = &\ \hat{w}\,(\nu)\,\frac{\partial w(x)}{\partial w\,(\nu)} + \hat{q}\,(\eta)\,\frac{\partial w(x)}{\partial q\,(\eta)} \\
&+\hat{B}\frac{\partial w(x)}{\partial \overline{w}}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{\phi}{(1-\phi)} - w\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}}\right] \\
&+\hat{\phi}\frac{\partial w(x)}{\partial \overline{w}}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{K}{(1-\phi)\phi} - w\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\}\right] \\
&+\hat{\rho}\left[\frac{\partial w(x)}{\partial \overline{w}}w\frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho} + \frac{\partial w(x)}{\partial \rho}\right].
\end{aligned}
\tag{B5}
$$

Where $q(\cdot)$ is the statistical density function of η, $p(\theta)$ is the statistical function of θ, $w(\cdot)$ is the density function of wealth x. $\hat{\cdot}$ is the perturbation of the corresponding variable $\cdot$. The stable distribution of wealth has five channels. The first term is the direct channel of x which describes the effect of changes in debt on wealth itself; The second term is the pertubation of the statistical density function of η, which describes the effect of the change of debt on the labor shock. Other channels include constant items of change in public debt $\hat{B}$, investment channels $\hat{\phi}$ and consumption channels $\hat{\rho}$.

Among them, we directly express some variable coefficients by the definition of their derivatives, for example, $\frac{\partial w(x)}{\partial \overline{w}}$ refers to the coefficient obtained after taking the derivative of the distribution $w(x)$ with respect to $\overline{w}$. The specific expression can be found in the section of the corresponding disturbance variables. (See the Appendix C: Equation (C3) for the derivative coefficients for K; Equation (C8) for the derivative coefficients for w).

B2. Young-age income distribution

Proof of the Law of Motion for Young-age Income

1. (y_{t+1}) solves $y_{t+1} = (1-\tau_{z,t+1})z_{t+1} + (1-\tau_{w,t+1})\eta_{t+1}w_{t+1}$

$y_{t+1} = h_{t+1}\varphi + (1-\tau_{w,t+1})\eta_{t+1}w_{t+1}$,

$h_{t+1} = y_t(1-\rho_t)\left\{\omega_{t+1}\phi_{t+1} + R_{t+1}(1-\phi_{t+1})\right\}$,

And, the dynamics of young-age income can be summarized as follows:

$$y_t : \quad y_{t+1} = y_t(1-\rho_t)J(\theta)\varphi + (1-\tau_{w,t+1})\eta_{t+1}w. \tag{B6}$$

Proof of the Stationary Distribution as a Fixed Point

1. $(G_{(y,t+1)}(y))$ solves $G_{(y,t+1)}(y) = P\left\{y_{t+1} \leq y\right\}$, the cumulative distribution function of y_{t+1}, $G_{(y,t+1)}(y)$, is:

$G_{(y,t+1)}(y) = P\left\{y_{t+1} \leq y\right\} = P\left\{y_t(1-\rho)\varphi J(\theta) + \eta\overline{w} \leq y\right\}$,

$G_{(y,t+1)}(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_{(y,t)}\left(\frac{y-\eta\overline{w}}{(1-\rho)\varphi J(\theta)}\right)p(\theta)q(\eta)d\theta d\eta$,

Let, $\mu = \frac{y-\eta\overline{w}}{(1-\rho)\varphi J(\theta)}$, so $\eta = \frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}$,

$$G_{(y,t+1)}(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} G_t(\mu)p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{(1-\rho)\varphi J(\theta)}{\overline{w}}d\theta d\mu \tag{B7}$$

2. $(g_{Y_{t+1}}(y))$ solves $\frac{\partial G_{(y,t+1)}(y)}{\partial y}$, for all y ≥ 0, the probability density function of

y_{t+1} is:

$$g_{Y_{t+1}}(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_{Y_t}\left(\frac{y-\eta\overline{w}}{(1-\rho)\varphi J(\theta)}\right)\frac{1}{(1-\rho)\varphi J(\theta)}p(\theta)q\left(\eta\right)d\theta d\eta, \tag{B8}$$

3. $(g(y))$ solves $g_{Y_t}(y) = g_{Y_{t+1}}(y)$ for all $y \geq 0$, the stationary distribution of $\{y_t\}_{t=0}^{\infty}$ satisfies

$$g_Y(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g_Y\left(\mu\right)p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{\overline{w}}d\theta d\mu, \tag{B9}$$

Proof of the Pertubation on Young-age Income The stable distribution of personal income obeys the probability density function $g(\cdot)$ and cumulative distribution function $G(\cdot)$. The effect of a debt change $\hat{B}$ of initial debt on the distribution of individual income, $\hat{g(y)}$, is the solution to the equation.

1. $(\hat{g(y)})$ solves $\dfrac{\partial g(y)}{\partial B}$ with Equation (C9), for all $\theta \in \Theta$

$$\hat{g(y)} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \hat{g}\left(\mu\right)p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{\overline{w}}d\theta d\mu$$

$$+\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g\left(\mu\right)p(\theta)\hat{q}\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{\overline{w}}d\theta d\mu$$

$$-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g\left(\mu\right)p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{(\overline{w})^2}\hat{B}\frac{\partial\hat{\overline{w}}}{\partial\hat{B}}d\theta d\mu$$

$$-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g\left(\mu\right)p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{(\overline{w})^2}\hat{\phi}\frac{\partial\hat{\overline{w}}}{\partial\hat{\phi}}d\theta d\mu$$

$$-\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g\left(\mu\right)p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{(\overline{w})^2}\hat{\rho}\frac{\partial\hat{\overline{w}}}{\partial\hat{\rho}}d\theta d\mu,$$

$$\begin{aligned}
g(\hat{y}) &= g\left(\hat{\mu}\right)\frac{\partial g(y)}{\partial g\left(\mu\right)} + \hat{q}\left(\eta\right)\frac{\partial g(y)}{\partial q\left(\eta\right)} \\
&+ \hat{B}\frac{\partial g(y)}{\partial \overline{w}}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{\phi}{(1-\phi)} - w\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}}\right] \\
&+ \hat{\phi}\frac{\partial g(y)}{\partial \overline{w}}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{K}{(1-\phi)\phi} - w\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\}\right] \\
&+ \hat{\rho}\frac{\partial g(y)}{\partial \overline{w}}\left[w\frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\right]
\end{aligned} \tag{B10}$$

Where $q(\cdot)$ is the statistical density function of η, $p(\theta)$ is the statistical function of θ, $g(\cdot)$ is the density function of y. $\hat{\cdot}$ is the perturbation of the corresponding variable $(\cdot)$. The stable distribution of young-age income has five channels. The first term is the direct channel of y which describes the effect of changes in debt on young-age income itself; The second term is the pertubation of the statistical density function of η, which describes the effect of the change of debt on the labor shock. Other channels include constant items of change in public debt $\hat{B}$, investment channels $\hat{\phi}$ and consumption channels $\hat{\rho}$.

Among them, we directly express some variable coefficients by the definition of their derivatives, for example, $\dfrac{\partial g(y)}{\partial \overline{w}}$ refers to the coefficient obtained after taking the derivative of the distribution $g(y)$ with respect to $\overline{w}$. The specific expression can be found in the section of the corresponding disturbance variables (See the Appendix C: Equation (C3) for the derivative coefficients for K; Equation (C8) for the derivative coefficients for w).

B3. Welfare distribution

Given that equation fully characterize the equilibrium for a given exemption level, it is now possible to understand how changes in the optimal level B affect social welfare.

Definition B.1 (Calculate a Single Integral of Welfare) *We get the optimal welfare function $(V(y))$ is given by*

1. $(V(y))$ solves $V(y) = \frac{\rho^{1-\gamma} + \beta\Gamma(1-\rho)^{1-\gamma} E\left\{J(\theta)^{1-\gamma}\right\}}{1-\gamma} \int y^{1-\gamma} g(y) dy$,

$y_i, i \in N$ *Generate N uniform numbers within* $[y_{min}, y_{max}]$,

Choose $y_1, ...y_N$,

We get the area as follows $S_1 = (y_{max} - y_{min})(V(y_1), S_2, ..., S_N$,

$\int y^{1-\gamma} g(y) dy = \lim_{N\to\infty} \frac{(y_{max} - y_{min})}{N} \sum_i y_i^{1-\gamma} g(y_i)$.

Because firm make zero profit in equilibrium, individuals maximizing indirect utility $V(B)$, defined in equation, maximizes social welfare in this economy. Lemma and proposition, which presents the central result of this paper, provides a test for whether to optimally increase or decrease the public debt. The effect of a debt reform $\hat{B}$ of the initial debt B on social welfare, $V\hat{(y)}$, is the solution to the functional equation. The perturbation formula for welfare is as follows,

Proof of the Channel Decomposition on Welfare

1. $(V\hat{(y)})$ solves $\frac{\partial V(y)}{\partial B}$ with $\left(\rho^{-\gamma} - \beta\Gamma(1-\rho)^{-\gamma} E\left\{J(\theta)^{1-\gamma}\right\}\right) = 0$ [3] and $J\hat{(\theta)}$ (Equation 22).

$$V\hat{(y)} = \frac{\beta\Gamma(1-\rho)^{1-\gamma}}{1-\gamma} E\left\{(1-\gamma)J(\theta)^{-\gamma}\hat{J}(\theta)\right\} \int y^{1-\gamma} g(y) dy$$
$$+ \frac{\rho^{1-\gamma} + \beta\Gamma(1-\rho)^{1-\gamma} E\left\{J(\theta)^{1-\gamma}\right\}}{1-\gamma} \int y^{1-\gamma} g\hat{(y)} dy,$$
$$\frac{\partial V(y)}{\partial EJ(\theta)} = \frac{\beta\Gamma(1-\rho)^{1-\gamma}}{1-\gamma} \int y^{1-\gamma} g(y) dy,$$
$$\frac{\partial V(y)}{\partial g(y)} = \frac{\rho^{1-\gamma} + \beta\Gamma(1-\rho)^{1-\gamma} E\left\{J(\theta)^{1-\gamma}\right\}}{1-\gamma} \int y^{1-\gamma} dy.$$

[3] According to the expression of ρ (See Equation (8))

$$V\hat{(y)} = \frac{\partial V(y)}{\partial EJ(\theta)} E\left\{(1-\gamma)J(\theta)^{-\gamma}\hat{J}(\theta)\right\} + \frac{\partial V(y)}{\partial g(y)} g\hat{(y)}.$$

(B11)

$$\begin{aligned}
V\hat{(y)} = {} & \hat{B}\frac{\partial V(y)}{\partial EJ(\theta)} E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\frac{\partial\omega(\theta)}{\partial K}\frac{\phi}{(1-\phi)}\phi + \frac{\phi}{\frac{\partial K}{\partial\tau_w}}\frac{\partial R}{\partial\tau_w}\right]\right\} \\
& + \hat{\phi}\frac{\partial V(y)}{\partial EJ(\theta)} E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\begin{array}{c}\frac{\partial\omega(\theta)}{\partial K}\frac{K}{(1-\phi)} + \omega(\theta) - R + \frac{\partial R}{\partial\phi}(1-\phi) \\ +\frac{\partial R}{\partial\tau_w}\frac{1}{\frac{\partial K}{\partial\tau_w}}\left(\frac{K}{\phi} - \frac{\partial K}{\partial\phi}(1-\phi)\right)\end{array}\right]\right\} \\
& + \hat{\rho}\frac{\partial V(y)}{\partial EJ(\theta)} E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\frac{\partial R}{\partial\rho} - \frac{1}{\frac{\partial K}{\partial\tau_w}}\frac{\partial K}{\partial\rho}\frac{\partial R}{\partial\tau_w}\right](1-\phi)\right\} \\
& + g\hat{(y)}\frac{\partial V(y)}{\partial g(y)}.
\end{aligned}$$

Proof of Proposition VI.1 (Sufficient statistics for optimal public in B)

1. (B^*) solves $V\hat{(y)} = 0$ with Equation (B11)

$$0 = \hat{B}\frac{\partial V(y)}{\partial EJ(\theta)} E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\frac{\partial\omega(\theta)}{\partial K}\frac{\phi}{(1-\phi)}\phi + \frac{\phi}{\frac{\partial K}{\partial\tau_w}}\frac{\partial R}{\partial\tau_w}\right]\right\}$$

$$+\hat{\phi}\frac{\partial V(y)}{\partial EJ(\theta)} E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\begin{array}{c}\frac{\partial\omega(\theta)}{\partial K}\frac{K}{(1-\phi)} + \omega(\theta) - R + \frac{\partial R}{\partial\phi}(1-\phi) \\ +\frac{\partial R}{\partial\tau_w}\frac{1}{\frac{\partial K}{\partial\tau_w}}\left(\frac{K}{\phi} - \frac{\partial K}{\partial\phi}(1-\phi)\right)\end{array}\right]\right\}$$

$$+\hat{\rho}\frac{\partial V(y)}{\partial EJ(\theta)} E\left\{(1-\gamma)J(\theta)^{-\gamma}\left[\frac{\partial R}{\partial\rho} - \frac{1}{\frac{\partial K}{\partial\tau_w}}\frac{\partial K}{\partial\rho}\frac{\partial R}{\partial\tau_w}\right](1-\phi)\right\} + g\hat{(y)}\frac{\partial V(y)}{\partial g(y)}.$$

The optimal public debt level means that (if public debt increase 1 unit, $\hat{B} = 1$;

if public debt decrease 1 unit, $\hat{B}=-1$), Let's go through the formula, we have:

(B12)

$$B^* = \frac{E\left\{J(\theta)^{-\gamma}\left[\begin{array}{c} -\hat{B}\underbrace{\left[K(\omega(\theta)-1+\delta)\right]}_{>0} \\ -\epsilon_{\phi,B}K^2\underbrace{\left[\frac{\partial\omega(\theta)}{\partial K}\frac{1}{(1-\phi)}+\frac{1}{\phi}\left(\omega(\theta)-R+\frac{\frac{\partial R}{\partial \tau_w}}{\frac{\partial K}{\partial \tau_w}}\right)\right]}_{<0}\end{array}\right]\right\}}{E\left\{J(\theta)^{-\gamma}\left[\begin{array}{c}\hat{B}\underbrace{K\frac{\frac{\partial R}{\partial \tau_w}}{\frac{\partial K}{\partial \tau_w}}}_{<0}+\epsilon_{\phi,B}\phi\underbrace{\left[\frac{\partial R}{\partial \phi}-\frac{\frac{\partial R}{\partial \tau_w}}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \phi}\right]}_{>0} \\ +\epsilon_{\rho,B}\rho\underbrace{\left[\frac{\partial R}{\partial \rho}-\frac{\frac{\partial R}{\partial \tau_w}}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\right]}_{<0}\end{array}\right]\right\}+\frac{\partial g(y)}{\partial B}\underbrace{\frac{\frac{\partial V(y)}{\partial g(y)}}{(1-\gamma)\frac{\partial V(y)}{\partial EJ(\theta)}}\frac{K}{\phi}}_{<0}}$$

B4. Theory for solve stable distributions

We introduce another method for solving stable distributions, using the loss function and optimizer of machine learning, to automatically find stable distributions. we have:

$$w(x)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}w_X(\nu)\,p(\theta)q\left(\frac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\frac{1}{(1-\rho)\overline{w}}d\theta d\nu,$$

$$g(y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}g_Y(\mu)\,p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{\overline{w}}d\theta d\mu,$$

Let, $T_x(\theta,\nu)_{(n,n)}=p(\theta)q\left(\frac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\frac{1}{(1-\rho)\overline{w}},$

$$T_y(\theta,\mu)_{(n,n)}=p(\theta)q\left(\frac{y-\mu(1-\rho)\varphi J(\theta)}{\overline{w}}\right)\frac{1}{\overline{w}}.$$

(B13)

$$\left[w(x)_1,...,w(x)_n\right]^T-\begin{bmatrix}\int_{-\infty}^{\infty}T_x(\theta,\nu)_1\left[w(x)_1,...,w(x)_n\right]^T \\ ... \\ \int_{-\infty}^{\infty}T_x(\theta,\nu)_n\left[w(x)_1,...,w(x)_n\right]^T\end{bmatrix}=\left[0,...,0\right]^T,$$

$$\left[g(y)_1,...,g(y)_n\right]^T-\begin{bmatrix}\int_{-\infty}^{\infty}T_y(\theta,\mu)_1\left[g(y)_1,...,g(y)_n\right]^T \\ ... \\ \int_{-\infty}^{\infty}T_y(\theta,\mu)_n\left[g(y)_1,...,g(y)_n\right]^T\end{bmatrix}=\left[0,...,0\right]^T.$$

Definition B.2 (Stable distribution of machine learning solutions) *Using the n equations of Equation (B13), we need to find a suitable result that satisfies the n equations simultaneously, $(w(x), g(y))$ are given by*

1. $(w(x))$ solves $F_x = w(x) - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(\nu)T_x(\theta,\nu)d\theta d\nu$,

$(g(y))$ solves $F_y = g(y) - \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(\mu)T_y(\theta,\mu)d\theta d\mu$.

2. Calculation procedure, taking wealth as an example,

(1) Create Learning Parameters: $F_{x,target} = \boldsymbol{0}, F_{x,pred} = F_x$,

(2) Neural network model (choose Loss function and optimizer),

$criterion = torch.nn.CrossEntropyLoss()$,

$optimizer = torch.optim.Adam(model.parameters(), lr = 1e-4)$.

By comparing the target data (F_{target}) with the forecast data (F_{pred}), the loss function can be constructed. In this paper, the common mean square error (MSE) is used to find the optimal combination of $w(x)$ to minimize the loss. Iterate $w(x)$ repeatedly to make the loss smaller and smaller. Similar to a blind person going down a mountain, standing on the mountain looking for the way down the mountain, given a random $w(x)$ initial point, looking for the fastest direction to travel. The gradient descent process, which is available in pytorch with the *backward()* command.

After understanding the learning parameters of the model, we need to build our own neural network. The neural network consists of forward propagation and backward propagation. Among them, the process of forward propagation refers to the transmission from the input layer to the hidden layer (if there are multiple hidden layers, propagation one by one), and from the hidden layer to the output layer. The backward propagation process means that we use the F_{pred} value and target value F_{target} calculated by forward propagation to obtain the loss function, and use the result of loss function to feed back from the output layer to the hidden layer and then to the input layer.

Definition B.3 (How to create neural network model) *we built a three-layer* *neural network structure, Input layer* $\underbrace{\rightarrow}_{n\ to\ n1}$ $\underbrace{\textit{Hidden layer}}_{Relu}$ $\underbrace{\rightarrow}_{n1\ to\ n}$ $\underbrace{\textit{Output layer}}_{Sigmoid}$*. We can have more than one hidden layer.*

1. Input layer: input the number of n neurons, the number of n neurons are transmitted to the hidden layer, and the number of n1 neurons are output.

2. Hidden layer: the Relu activation function is passed, the number of n1 neurons enter the output layer.

3. Output layer: the number of n neurons are output and activated by the Sigmoid function.

Although the whole neuron is complex, it has the same local characteristics and is still a linear relationship, $z = \sum w_i x_i + b$, and an activation function $\sigma(z)$. In machine learning, there are many commands about activation functions. In our paper, we mainly use *Relu*[4] and *Sigmoid* [5] activation functions. Both types of activation functions belong to the category of nonlinear activation.

Proposition B.1 (The processes of forward propagation) *In our paper, only the use of one hidden layer is involved. The upper corner is the network layer, and the lower corner is the index of the neuron. For example, w_{23}^2, 2 represents the second layer neuron, and 23 represents the second neuron in the third layer to the third neuron in the second layer.*

Input layer: n neurons, from g_1 to g_n,

Hidden layer1: sz neurons, from a_1 to a_{sz},

Output layer: 1 neurons, G_1.

1. From input layer to hidden1 layer:

$a_1^2 = \sigma(z_1^2) = \sigma(w_{11}^2 x_1 + w_{12}^2 x_2 + ... + w_{1n}^2 x_n + b_1)$,

...,

$a_{sz}^2 = \sigma(z_{sz}^2) = \sigma(w_{sz,1}^2 x_1 + w_{sz,2}^2 x_2 + ... + w_{sz,n}^2 x_n + b_{sz})$.

2. From hidden1 layer to hidden2 layer:

[4]ReLU: This activation function is to take all negative values to zero and leave positive values unchanged. $f(x) = max(0, x)$

[5]Sigmoid: This activation function takes a real number as input and a numeric value between 0 and 1 as output. $\sigma(x) = 1/(1 + e^{-x})$

$G_1^3 = \sigma(z_1^3) = \sigma(w_{1,1}^3 a_1 + w_{1,2}^3 a_2 + ... + w_{1,sz}^3 a_{sz} + a_1^3)$.

The results of the upper and lower subscripts of ***bias****, active function and* ***output*** *are consistent. The general form as follow,* $a_j^L = \sigma(z_j^L) = \sigma\left(\sum_{k=1}^m w_{jk}^L a_k^{L-1} + b_j^L\right)$.

Form of vector, we have: $\mathbf{a^L}$ *has m neurons with dimension (m*1);* $\mathbf{a^{L-1}}$ *has n neurons with dimension (n*1);* $\mathbf{w^L}$ *with dimension (m*n);* $\sigma(\mathbf{z^L})$ *with dimension (m*1); And* $\mathbf{b^L}$ *with dimension (m*1),* $\mathbf{a^L} = \sigma(\mathbf{z^L}) = \sigma\left(\mathbf{w^L a^{L-1}} + \mathbf{b^L}\right)$.

Definition B.4 (Calculate the Dual Integral of Stable Distribution) *We get the stable distribution of wealth* $(w(x))$ *is given by*

$w(x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(\nu)\, q\left(\frac{x - \nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\frac{1}{(1-\rho)\overline{w}} d\nu p(\theta) d\theta$, *the integral result is obtained by calculating the volume of the curved top column in B1(1).*

$x_i, \nu_k, i \in N, k \in N$ *Go to the uniformly distributed N points in* $[x_{min}, x_{max}]$,

$\theta_j, j \in N$ *Generate N uniformly distributed numbers, where* $log(\theta) \sim N(\mu_\theta, \sigma_\theta^2)$. *Fixing an Income and Shock* y_i, θ_j, *get the red area* $(h(\theta_j))$ *in Figure B1(2)(3),*

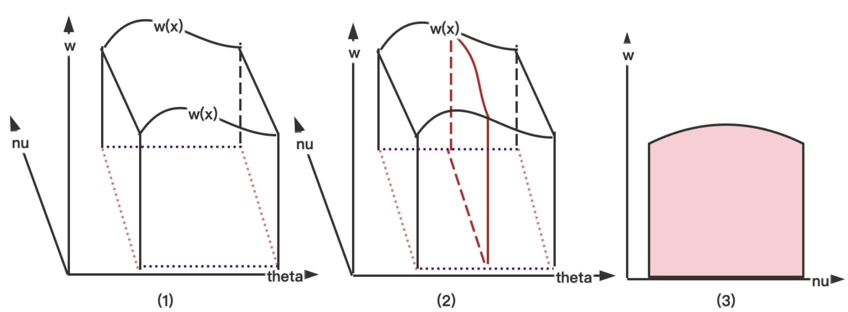

Figure B1. Monte carlo Dual Integral

take multiple random numbers in the range of θ, *get multiple section data, then get average and get approximation of* $(h(\theta))$

1.$(h(\theta_j))$ *solves*

$h(\theta_j) = \frac{x_{max} - x_{min}}{N}\sum_{k=0}^{k=N} w(\nu_k)\, q\left(\frac{x_i - \nu_k(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\frac{1}{(1-\rho)\overline{w}}$,

2.$(h(\theta))$ *solves* $h(\theta) = \frac{1}{N}\sum_{j=0}^{j=N} h(\theta_j)$,

3.$(w(x_i))$ *solves*

$w(x_i) = \int_{-\infty}^{\infty} h(\theta)p(\theta)d\theta = \frac{\theta_{max} - \theta_{min}}{N} \sum_{j=0}^{j=N} h(\theta_j)p(\theta_i) = h(\theta)$,

4.$(w(x))$ solves $[w(x_1), ..., w(x_N)]^T$.

5. ***Algorithm see Table D2***

Technical appendix for comparative static analysis

We will give a technical appendix for solving nonlinear equations using comparative static analysis methods. In our paper, we mainly solve Equations (7) and (8). We end up with our perturbation term of the endogenous variable.

Proof of channel decomposition of variable(comparative static analysis)

1. (R) Simplify the Equation (A16) with $\lambda, \lambda_i, i = 1, 2$

$$\lambda = \frac{\varphi\alpha(1-\tau_p) + (1-\tau_w)(1-\alpha)}{\varphi\alpha + (1-\tau_w)(1-\alpha)\left(1 - \varphi\frac{\tau_z}{1-\tau_z}\right) + \varphi(1-\alpha)\tau_w}$$

$$\lambda_1 = \frac{1}{\varphi\alpha(1-\tau_p) + (1-\tau_w)(1-\alpha)}\left((1-\alpha)\tau_w + \frac{\tau_z\varphi\alpha(1-\tau_p)}{1-\tau_z} + \tau_p\alpha\right)$$

$$\lambda_2 = \frac{(1-\delta)\varphi}{\varphi\alpha(1-\tau_p) + (1-\tau_w)(1-\alpha)}\left(-(1-\alpha)\tau_w\frac{1}{1-\tau_z} + \frac{\tau_z(1-\alpha)}{1-\tau_z} - \tau_p\alpha\right)$$

$$R = \lambda\left[1 + \lambda_1\frac{1}{(1-\phi)(1-\rho)} + \lambda_2\frac{\phi}{1-\phi}\right]. \tag{C1}$$

2. $(\hat{R})$ solves $\frac{\partial R}{\partial B}$ with Equation (A16) and (C1)

$$\frac{\partial R}{\partial \phi} = \lambda\left[\frac{\lambda_1}{(1-\phi)^2(1-\rho)} + \frac{\lambda_2}{(1-\phi)^2}\right],$$

$$\frac{\partial R}{\partial \rho} = \frac{\lambda\lambda_1}{(1-\phi)(1-\rho)^2},$$

$$
\frac{\partial R}{\partial \tau_w} = \lambda \left\{ \begin{array}{l}
\frac{1}{(1-\phi)(1-\rho)} \left\{ \begin{array}{c} \frac{(1-\alpha)\left((1-\alpha)\tau_w + \frac{\tau_z \varphi\alpha(1-\tau_p)}{1-\tau_z} + \tau_p\alpha\right)}{[\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)]^2} \\ + \frac{1-\alpha}{\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)} \end{array} \right\} \\
+\frac{\phi}{1-\phi} \left\{ \begin{array}{c} \frac{(1-\delta)\varphi(1-\alpha)\left(\frac{-(1-\alpha)\tau_w}{1-\tau_z} + \frac{\tau_z(1-\alpha)}{1-\tau_z} - \tau_p\alpha\right)}{[\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)]^2} \\ - \frac{(1-\delta)\varphi(1-\alpha)}{\varphi\alpha(1-\tau_p)(1-\tau_w)(1-\alpha)} \frac{1}{1-\tau_z} \end{array} \right\} \\
-R \left\{ \begin{array}{c} \frac{\varphi(1-\alpha)\left[(1-\alpha)\tau_w \frac{1}{1-\tau_z} + \tau_p\alpha - \frac{\tau_z}{1-\tau_z}(1-\alpha)\right]}{[\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)]^2} \\ + \frac{\varphi(1-\alpha)}{\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)} \frac{1}{1-\tau_z} \end{array} \right\}
\end{array} \right\},
$$

(C2)
$$
\hat{R} = \hat{\phi}\frac{\partial R}{\partial \phi} + \hat{\rho}\frac{\partial R}{\partial \rho} + \hat{\tau_w}\frac{\partial R}{\partial \tau_w}
$$

3. $(\hat{K})$ solves $\frac{\partial K}{\partial B}$ with K (Equation A15)

$$
\frac{\partial K}{\partial \phi} = \frac{R\frac{\varphi}{\phi^2} - \frac{1}{\phi^2(1-\rho)} - \frac{(1-\phi)\varphi}{\phi}\frac{\partial R}{\partial \phi}}{(\alpha-1)K^{\alpha-2}A\left(E\theta_t^{\frac{1}{\alpha}}\right)^{\alpha}\{\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)\}},
$$

$$
\frac{\partial K}{\partial \rho} = \frac{\left(\frac{1}{\phi(1-\rho)^2} - \frac{(1-\phi)\varphi}{\phi}\frac{\partial R}{\partial \rho}\right)}{(\alpha-1)K^{\alpha-2}A\left(E\theta_t^{\frac{1}{\alpha}}\right)^{\alpha}\{\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)\}},
$$

$$
\frac{\partial K}{\partial \tau_w} = \left\{ \begin{array}{c}
-\frac{\frac{1}{\phi(1-\rho)} - \varphi(1-\delta) - \frac{(1-\phi)\varphi R}{\phi}}{A\left(E\theta_t^{\frac{1}{\alpha}}\right)^{\alpha}K^{\alpha-2}\{\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)\}^2} \\
-\frac{\frac{(1-\phi)\varphi}{\phi}\frac{\partial R}{\partial \tau_w}}{(\alpha-1)K^{\alpha-2}A\left(E\theta_t^{\frac{1}{\alpha}}\right)^{\alpha}\{\varphi\alpha(1-\tau_p)+(1-\tau_w)(1-\alpha)\}}
\end{array} \right\}
$$

(C3)
$$
\hat{K} = \hat{\phi}\frac{\partial K}{\partial \phi} + \hat{\rho}\frac{\partial K}{\partial \rho} + \hat{\tau_w}\frac{\partial K}{\partial \tau_w}
$$

4. $(\hat{\tau_w})$ solves $B = (\frac{1}{\phi} - 1)K$

$$\hat{B} = (\frac{1}{\phi} - 1)\hat{K} - \hat{\phi}\frac{K}{\phi^2} = (\frac{1}{\phi} - 1)\left[\hat{\phi}\frac{\partial K}{\partial \phi} + \hat{\rho}\frac{\partial K}{\partial \rho} + \hat{\tau_w}\frac{\partial K}{\partial \tau_w}\right] - \hat{\phi}\frac{K}{\phi^2}$$

$$\hat{\tau_w} = \hat{B}\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}} + \hat{\phi}\frac{1}{\dfrac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\} - \hat{\rho}\frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}. \tag{C4}$$

Rewrite the perturbed expressions for R and K, we have:

$$\begin{aligned} \hat{R} = {} & \hat{B}\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}}\frac{\partial R}{\partial \tau_w} + \hat{\phi}\left[\frac{\partial R}{\partial \phi} + \frac{\partial R}{\partial \tau_w}\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\}\right] \\ & + \hat{\rho}\left[\frac{\partial R}{\partial \rho} - \frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\frac{\partial R}{\partial \tau_w}\right]. \end{aligned} \tag{C5}$$

$$\hat{K} = \hat{B}\frac{\phi}{(1-\phi)} + \hat{\phi}\frac{K}{(1-\phi)\phi}. \tag{C6}$$

5. $(\hat{\omega(\theta)})$ solves $\dfrac{\partial \omega(\theta)}{\partial B}$ with $\hat{K}$ (Equation C6)

$$\omega(\theta) = \alpha A(1-\tau_p)\theta^{\frac{1}{\alpha}}\left(E\theta^{\frac{1}{\alpha}}\right)^{\alpha-1}K^{\alpha-1} + 1 - \delta,$$

$$\frac{\partial \omega(\theta)}{\partial K} = \alpha A(1-\tau_p)\theta^{\frac{1}{\alpha}}\left(E\theta^{\frac{1}{\alpha}}\right)^{\alpha-1}(\alpha-1)K^{\alpha-2},$$

$$\hat{\omega(\theta)} = \frac{\partial \omega(\theta)}{\partial K}\hat{K} = \hat{B}\frac{\partial \omega(\theta)}{\partial K}\frac{\phi}{(1-\phi)} + \hat{\phi}\frac{\partial \omega(\theta)}{\partial K}\frac{K}{(1-\phi)\phi}. \tag{C7}$$

6. $(\hat{w})$ solves $\dfrac{\partial w}{\partial B}$ with $\hat{K}$ (Equation C6)

$$\frac{\partial w}{\partial K} = (1-\alpha)A\left(E\theta_{t+1}^{\frac{1}{\alpha}}\right)^{\alpha}\alpha K^{\alpha-1},$$

$$\hat{w} = \hat{B}\frac{\partial w}{\partial K}\frac{\phi}{(1-\phi)} + \hat{\phi}\frac{\partial w}{\partial K}\frac{K}{(1-\phi)\phi} \tag{C8}$$

7. $(\hat{\overline{w}})$ solves $\dfrac{\partial \overline{w}}{\partial B}$ with $\hat{w}$ (Equation C8) and $\hat{\tau_w}$ (Equation C4)
$\hat{\overline{w}} = (1-\tau_w)\hat{w} - \hat{\tau_w} w$,

(C9)
$$\hat{\overline{w}} = \hat{B}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{\phi}{(1-\phi)} - w\frac{\phi}{(1-\phi)\frac{\partial K}{\partial \tau_w}}\right] + \hat{\rho}\left[w\frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\right]$$
$$+\hat{\phi}\left[(1-\tau_w)\frac{\partial w}{\partial K}\frac{K}{(1-\phi)\phi} - w\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{(1-\phi)\phi} - \frac{\partial K}{\partial \phi}\right\}\right].$$

8. $(J(\hat{\theta}))$ solves $\dfrac{\partial J(\theta)}{\partial B}$ and $J(\theta) = \omega(\theta)\phi + R(1-\phi)$
$J(\hat{\theta}) = \omega(\hat{\theta})\phi + \hat{\phi}\omega(\theta) + \hat{R}(1-\phi) - \hat{\phi}R = \omega(\hat{\theta})\phi + \hat{\phi}(\omega(\theta)-R) + \hat{R}(1-\phi)$,

(C10)
$$J(\hat{\theta}) = \hat{B}\left[\frac{\partial \omega(\theta)}{\partial K}\frac{\phi}{(1-\phi)}\phi + \frac{\phi}{\frac{\partial K}{\partial \tau_w}}\frac{\partial R}{\partial \tau_w}\right]$$
$$+\hat{\phi}\left[\frac{\partial \omega(\theta)}{\partial K}\frac{K}{(1-\phi)} + \omega(\theta) - R + \frac{\partial R}{\partial \phi}(1-\phi) + \frac{\partial R}{\partial \tau_w}\frac{1}{\frac{\partial K}{\partial \tau_w}}\left\{\frac{K}{\phi} - \frac{\partial K}{\partial \phi}(1-\phi)\right\}\right]$$
$$+\hat{\rho}\left[\frac{\partial R}{\partial \rho} - \frac{1}{\frac{\partial K}{\partial \tau_w}}\frac{\partial K}{\partial \rho}\frac{\partial R}{\partial \tau_w}\right](1-\phi).$$

9. **Algorithm see Table D3**

Technical appendix for algorithm

Table D1—: Numerical method to general equilibrium

Algorithm 1 Numerical method to general equilibrium
Input: parameters
$\beta, \delta, \alpha, \gamma, \tau_p, \tau_z, A, \chi, \varphi$
Continued on next page

Table D1– continued from previous page

$\left(\begin{array}{c}\textit{Parameters about subject discount factor, capital depreciation rate,}\\ \textit{capital share, risk aversion coefficient, profit tax, bequest tax, TFP,}\\ \textit{bequest motive coefficient and the after-tax inheritance income ratio}\end{array}\right)$

$(\mu_\theta, \sigma_\theta^2), (\mu_\eta, \sigma_\eta^2)$ (Random distribution of productivity and labor shocks)

$fun\_\omega, fun\_K, fun\_R, fun\_\rho, fun\_\phi$ (Equation in (A6, 11, 12, 8, 7))

$\tau_w, \tau_{wl}, \tau_{wr}$ (wage rate, minimum tax rate, maximum tax rate)

$dist\_B, dist\_fun$(Errors in the supply and demand of debt and equations)

$step_\rho, step_\phi$ (Influence the step of ρ and ϕ parameter updating)

Output: stable variables $[K, R, \rho, \phi, \tau_w]$

while $dist\_B > 1.0e-5$

$\tau_w = 0.5 * (\tau_{wl} + \tau_{wr}), step_\rho, step_\phi$ (Initialization values)

while $(dist\_fun > 1.0e-5)$ (Solving equation yields variables)

$R = fun\_R(\phi\_try, \rho\_try)$ (step 1 solve R and K)

$K = fun\_K(\phi\_try, \rho\_try)$

$\rho\_new = \mathbf{fsolve}(fun\_\rho, 0)$ (step 2 solve ρ)

$\phi\_new = \mathbf{fsolve}(fun\_\phi, 0)$ (step 3 solve ϕ)

$dist\_\phi, dist\_\rho, dist\_fun = \mathbf{max}([dist\_\rho, dist\_\phi])$(compute the errors)

ϕ_try, ρ_try (update by $step_\rho, step_\phi$)

$dist\_B$ (compute the errors of debt)

if $dist\_B > 0$ **then** $\tau_{wl} = \tau_w$ **else** $\tau_{wr} = \tau_w$

Table D2—: Machine learning to solve stable distributions

Algorithm 2 Machine learning to solve stable distributions (Wealth)

Input: For parameter settings see Algorithm 1 (Table D1)

F_target (The aim function)

Continued on next page

Table D2– continued from previous page

vec_x (The vector of wealth)

epoch (The number of learning when loss has never been achieved)

$dist\_acc$(Errors in the supply and demand of debt and equations)

Building a neural network: See Definition B.3

$n1$(Number of neurons in the hidden layer)

lr(Learning rate: influence the step of parameter updating)

$decision\_rule$(Building neutral networks according to $n1$)

$optimizer =$ **optim.Adam**($decision\_rule.parameters(), lr$)(Select Optimizer)

$criterion =$ **torch.nn.MSELoss**($reduction =' mean'$)(loss function)

Output: stable distribution about wealth $[w(x)]$

when $dist\_acc > 0$

$w(x) = decision\_rule(vec_x)$

$(Results\ of\ learning\ used\ in\ calculate\ vec_x\ by\ 'decision\_rule')$

$F\_pred \leftarrow \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} w(\nu)\, q\left(\frac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right)\frac{1}{(1-\rho)\overline{w}} d\nu p(\theta) d\theta$

(*Compute the 'F_pred' by **Definition**B.4*)

loss=criterion $\leftarrow 1/N\sum_{i=1}^{N}|F\_pred - F\_target|^2$

optimizer.zero_grad() (Clear the gradients calculated before)

loss.backward() $\leftarrow w(x)_{t+1} = w(x)_t - lr\frac{\partial L}{\partial w(x)}$ (update gradients)

optimizer.step() (update parameters in neutral network by optimizer)

The header of each sub-table has the number of iterations *Epoch* and *loss* that mark machine learning. The stable distribution results of machine learning and corresponding loss diagrams will be shown as follows (See form Table D1 to D4).

Table D3—: Machine learning to channel decomposition

Continued on next page

Table D3– continued from previous page

Algorithm 3 Machine learning to channel decomposition

Input: For parameter settings see Algorithm 1 and 2 (Table D1, D2)

vec_x, vec_share_x

(Vector of wealth before and after perturbation using interpolation)

wx, wx_share(PDF before and after perturbation using interpolation)

Output1: Channel decomposition of variable, Table[6, 7]

$variable =$ **torch.tensor**($variable$,**requires_grad=True**)

(Determine the variables to be derived)

$fun_{variable}$**.backward()** (Calculate the derivative of $R, K, \tau_w, w, \overline{w}, \omega, J$)

$grad_{variable} = variable$**.grad**(Calculate the gradient of a variable)

$hat_{variable} \leftarrow (variable_{t+1} - variable_t)/0.01$

(Calculate the derivative of $\hat{R}, \hat{\phi}, \hat{\rho}, \hat{K}, \hat{w}, \hat{\overline{w}}, \hat{\omega}, \hat{J}$)

$variable$**.grad.zero_()**(Zero gradient is required before the next derivation)

$channel_{variable} = hat_{variable} * grad_{variable}$

Output2: Channel decomposition of stable distribution $w(x)$

$hat\_wx_{variable}$(Calculate the perturbation results using the interpolated PDF)

def $Input\_x\_Output\_wx(want\_x\_vec)$

(Arbitrary input '$want\_vec_x$', output the corresponding '$output_{wx}$' value)

for $input\_x$ **in** $want\_vec_x$

for j **in range** ((N-1))

$$output+= \left(\begin{array}{c} \textbf{np.poly1d} \left(\textbf{np.polyfit} \begin{pmatrix} vec\_share_x[j:j+2] \\ , wx\_share[j:j+2] \end{pmatrix} \right) \\ (input_x)(vec\_share_x[j+1] > input_x >= vec\_share_x[j]) \end{array} \right)$$

$output_{wx} = output$ (Save in each pdf of $input\_x$)

return output

def $Input\_x\_Output\_derive(input\_x\_vec, input\_x, input\_wx)$

(Derivative gradient of the input matrix)

$variable =$ **torch.tensor**($variable$,**requires_grad=True**)

Continued on next page

Table D3– continued from previous page

$f \leftarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} w(\nu)\, q\left(\frac{x-\nu(1-\rho)\varphi J(\theta)}{(1-\rho)\overline{w}}\right) \frac{1}{(1-\rho)\overline{w}} d\nu p(\theta) d\theta$

f**.backward()**(Derive the equation f)

$grad\_wx_{variable} = variable$**.grad**(Derivative results of the variables)

return $grad\_wx_{wx}, grad\_wx_{PQ}, grad\_wx_{\rho}, grad\_wx_{\overline{w}}$

$output_{wx} = Input\_x\_Output\_wx(vec\_share_x)$

$result = Input\_x\_Output\_derive(vec\_share_x, input_x, output_{wx})$

(taking each value of the matrix $'vec\_share'_x$ as an $'input'_x$)

$channel\_wx_{variable} = grad\_wx_{variable} * hat\_wx_{variable}$

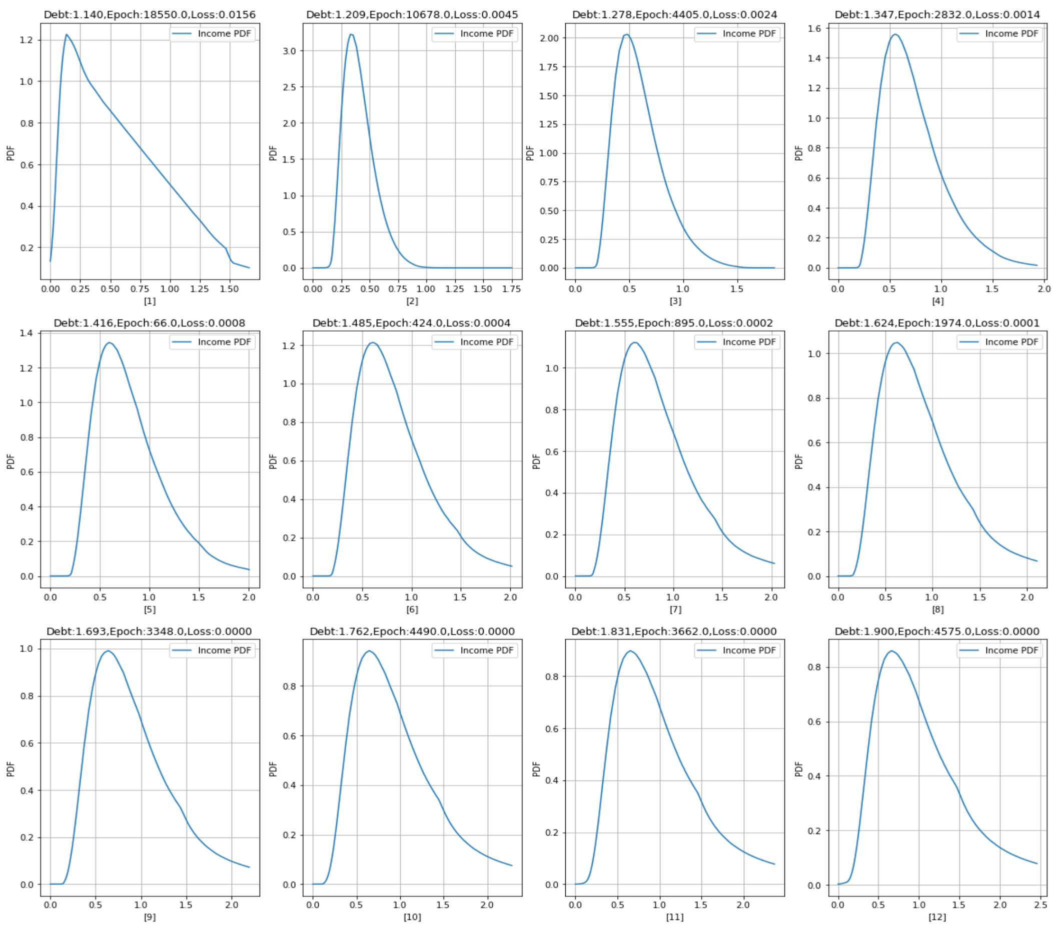

Figure D1. Stable distribution of young-age income

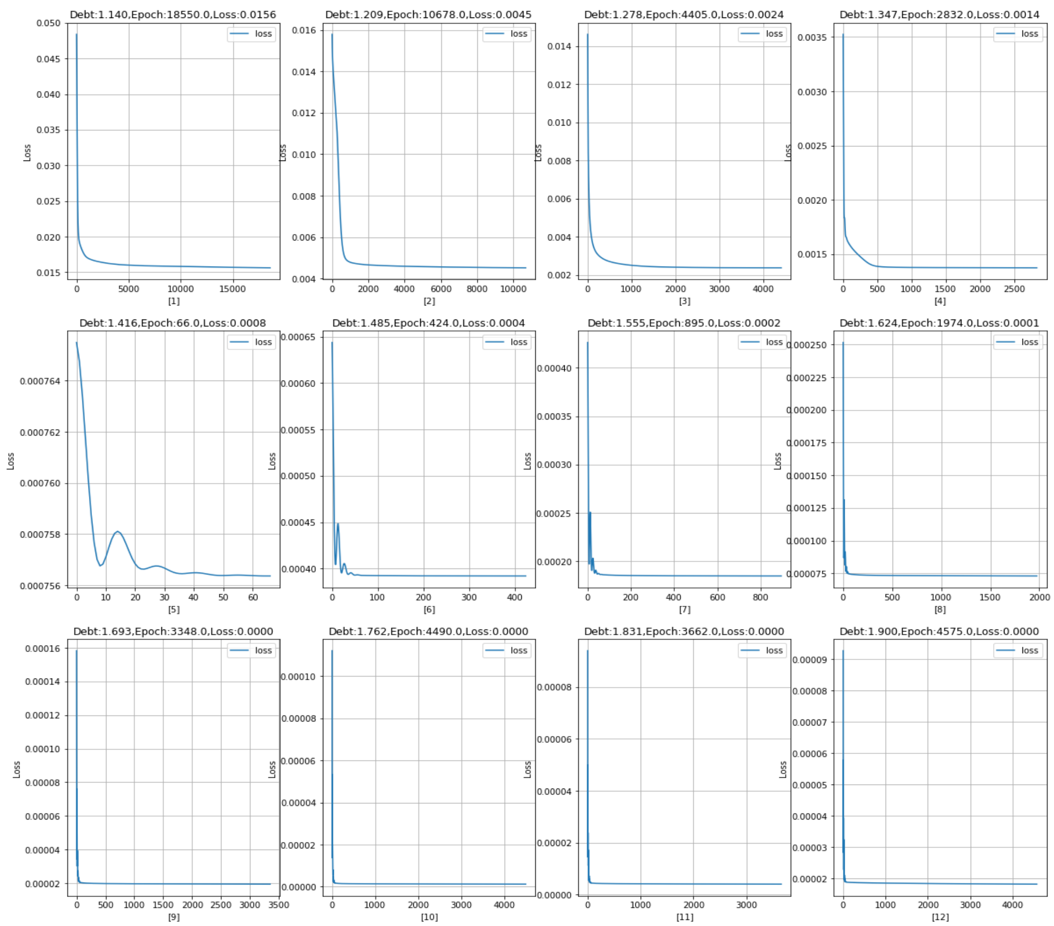

Figure D2. Loss of income

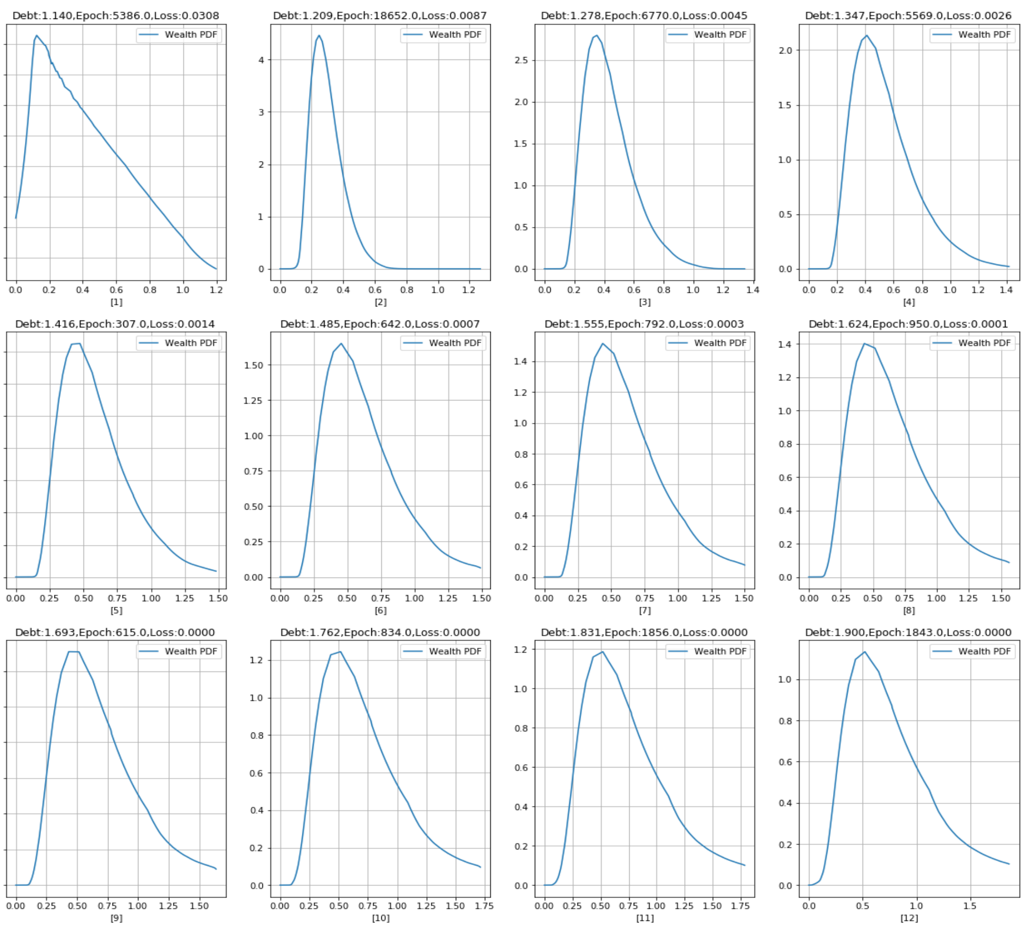

Figure D3. Stable distribution of wealth

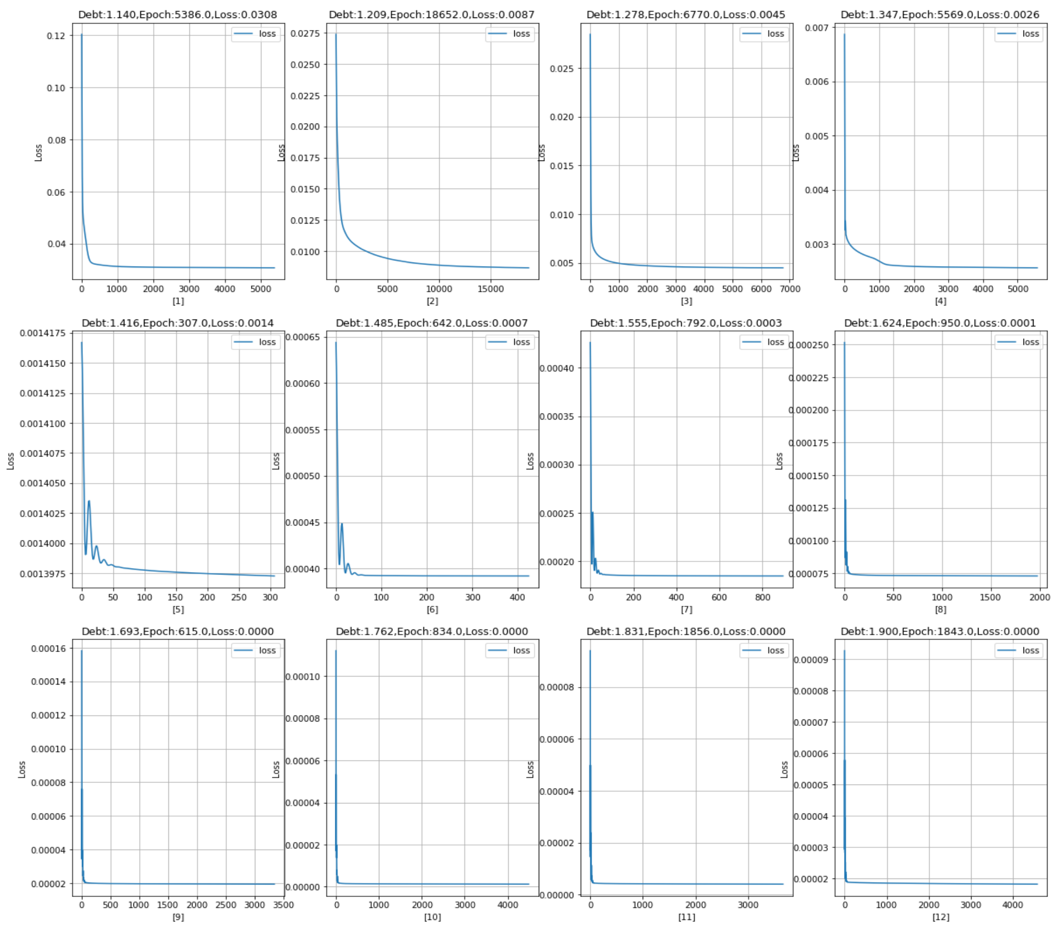

Figure D4. Loss of wealth