Capital Return Jumps and Wealth Distribution†

Jess Benhabib‡ Wei Cui§ Jianjun Miao¶

November 29, 2021

Abstract

The distributions of wealth in the US and many other countries are strikingly concentrated on the top and skewed to the right. To explain the income and wealth inequality, we provide a tractable heterogeneous-agent model with incomplete markets in continuous time. We separate illiquid capital assets from liquid bond assets and introduce capital return jump risks. Under recursive utility, we derive optimal consumption and wealth in closed form and show that the stationary wealth distribution has an exponential right tail. Our calibrated model can match the income and wealth distributions in the US data including the extreme right tail. We also study the effect of taxes on the distribution of wealth.

Keywords: Wealth distribution; Inequality; Heterogeneous agents; Incomplete markets; Exponential tail

JEL Classifications: C61, D83, E21, E22, E31.

†We thank Steve Kou, Franck Portier, Morten Ravn, and Hao Xing for helpful discussions.
‡Dept of Economics, New York University, Email: jess.benhabib@nyu.edu
§Dept of Economics, University College London, Email: w.cui@ucl.ac.uk
¶Dept of Economics, Boston University, Email: miaoj@bu.edu

1 Introduction

The distributions of wealth in the US and many other countries are strikingly concentrated on the top and skewed to the right (Piketty (2014), Saez and Zucman (2016), Bricker et al. (2016), and Smith, Zidar, and Zwick (2021)). For example, using the US administrative tax data, Smith, Zidar, and Zwick (2021) estimate that the top 0.1% and 1% wealth shares increased from 9.9% and 23.9% in 1989 to 15% and 31.5% in 2016, respectively. Understanding the sources of the wealth inequality and the mechanism that generates such inequality is important not only for policy makers, but also for academic researchers.

The goal of our paper is to provide a tractable model that accounts for the US distributions of earnings and wealth. Our model builds on the standard quantitative theory used in the heterogeneous-agent literature within macroeconomics: the Bewley-Huggett-Aiyagari (BHA) model (Bewley (1980), Huggett (1993), and Aiyagari (1994)). As is well known (e.g., Benhabib and Bisin (2018) and Stachurski and Toda (2019)), a standard BHA model with infinitely-lived agents facing idiosyncratic labor income risks alone generates a counterfactual result that the tail thickness of the model output (wealth) cannot exceed that of the input (income).

We depart from the standard BHA model by introducing two key ingredients. First, we introduce portfolio heterogeneity by separating illiquid capital assets from liquid safe assets (bonds). In the standard BHA model, both types of assets are perfect substitutes and earn the same constant return (interest rate) in a stationary equilibrium. In our model, capital assets are illiquid and incur adjustment costs (Kaplan and Violante (2004) and Kaplan, Moll, and Violante (2018)). Thus the capital return differs from the interest rate.

Second, we introduce idiosyncratic investment risks in the form of Poisson jumps of capital returns, which apply only to new capital investments, but not to rate of return on capital already in place. At each point in time, each household has a chance of conducting innovations/R&D. Such activities arrive as rare events and may generate large stochastic returns. These returns are critical to account for the top wealth shares. This feature is consistent with the wealth accumulation of some richest Americans in recent years. By examining 100 of them listed in the Forbes magazine, Graham (2021) argues that "[b]y 2020 the biggest source of new wealth was what are sometimes called 'tech' companies. Of the 73 new fortunes, about 30 derive from such companies. These are particularly common among the richest of the rich: 8 of the top 10 fortunes in 2020 were new fortunes of this type."

Incorporating the above two ingredients in a tractable continuous-time model, we make a technical contribution by adopting the affine-jump diffusion (AJD) framework of Duffie, Pan, and Singleton (2000) in the finance literature. Specifically, we assume that labor income follows a square-root process (Cox, Ingersol, and Ross (1985)) and each household's preferences are represented by continuous-time recursive utility of Duffie and Epstein (1992). We embed the power-exponential specification of the discrete-time recursive utility model of Weil (1993) into our continuous-time setup. This specification features a constant elasticity of intertemporal substitution (EIS) and a constant coefficient of absolute risk aversion (CARA). Abstracting away from binding borrowing constraints, we are able to derive a close-form solution to the individual consumption/saving problem under uninsurable Brownian labor income risk and Poisson capital return jump risk. Unlike the usual exponential-affine model (e.g, Caballero (1990), Angeletos and Calvet (2006), and Wang (2007)), our model setup delivers positive labor income and positive optimal individual consumption under some mild assumptions. The separation between EIS and CARA in our utility model is important not only for understanding precautionary saving (Weil (1993)), but also for generating a large marginal propensity to consume (MPC) as in the data.[1] This feature is critical for the existence of a stationary equilibrium and also for matching the data.

We provide three major theoretical results. First, we prove the existence of a stationary equilibrium in which the interest rate is lower than the subjective discount rate as in Aiyagari (1994). We show that the equilibrium prices and aggregate quantities can be determined independently of the full wealth distribution because only the mean matters for the aggregate variables. After the equilibrium prices are pinned down, our explicit solution for the optimal consumption and wealth processes allows us to simulate the wealth distribution tractably and efficiently.[2]

Second, we show that the joint equilibrium wealth and labor income process is an AJD process. Extending the method of Wang (2007), we provide an explicit recursive formula to compute the moments of the stationary wealth and labor income distributions. Our explicit formula shows clearly how the capital return jump intensity and the jump size distribution can generate a larger skewness and a larger kurtosis for the wealth distribution relative to the labor income distribution.

Because skewness and kurtosis may not fully characterize the tail behavior of a distribution, our third result, which is also the most novel theoretical result in the paper, studies this issue. We show that both the stationary wealth and labor income distributions have an exponential tail when the jump size follows a hyper-exponential distribution (HED), i.e., a finite mixture of exponential distributions (Feldmann and Whitt (1998)). Moreover, we explicitly characterize their exponential decay rates. We identify conditions on the HED such that the tail of the wealth distribution decays more slowly than that of the labor income distribution. In this case the wealth distribution has a heavier tail than the labor income distribution.

To examine the quantitative implications of our theory, we calibrate our model to confront the US data. We choose parameter values to match the US micro and macro data, and especially statistics related to the wealth and earnings distributions. Our adoption of the HED specification for the jump size allows us to compute the stationary equilibrium tractably because the HED delivers explicit formulas for the moment generating function and the Laplace transform. Moreover, the HED can approximate any completely monotone distribution including the Pareto distribution (Feldmann and Whitt (1998)) and thus is very flexible for us to match the wealth distribution data.[3]

[1] See Kaplan and Violante (2021) for discussions of the impact of recursive utility on the MPC in the discrete-time BHA framework.

[2] Because we have a closed-form solution for the wealth process, we do not need to use the PDE approach of Achdou et al. (2020) to solve for the wealth distribution. As Gouin-Bonenfant and Toda (2020) argue, the usual numerical methods may suffer from large truncation errors at the upper tail of the wealth distribution.

[3] A distribution with the pdf $g(x)$ is completely monotone if the n^{th} derivative $g^{(n)}(x)$ exists and $(-1)^n g^{(n)}(x) \geq 0$ for any $n \geq 1$.

By specifying three exponential components for the HED, we find that our calibrated model can match the wealth distribution in the data closely, including the wealth shares on the very top. In particular, we can match the wealth shares held by the top 0.1% and 1%. As is well known, matching such data in general equilibrium is a challenging task. The recent literature has focused on generating a Pareto tailed wealth distribution (e.g., Benhabib, Bisin, and Zhu (2011) and Gabaix et al. (2016)). By contrast, our quantitative results show that even an exponential tailed distribution can also match the data quite well. The reason is that it is very difficult to distinguish empirically the exponential-type tails from power-type (Pareto) tails given a finite sample of data (Heyde and Kou (2004)). Heyde and Kou (2004) show that sample sizes typically in the tens of thousands or even hundreds of thousands are necessary to distinguish power-type tails from exponential-type tails.

A nice feature of our model is that it can be applied to study the impact of taxing the super rich. Specifically, consider taxation on the capital return jumps drawn from the component distribution of the HED with the largest mean. We find that the impact on wealth inequality depends crucially on how tax revenues are distributed. When tax revenues are transferred to all households evenly, such a policy raises wealth inequality. But when tax revenues are used to inject credit to borrowers, such a policy reduces wealth inequality.

Related literature. Our paper contributes to the macroeconomics literature on wealth inequality in the tradition of the BHA model.[4] Early studies have tried to modify assumptions in the BHA model to make wealthier agents want to save more. Examples include the warm-glow bequest and human capital motives of De Nardi (2004), very large earnings risk for high-earning households of Castaneda, Diaz-Gimenez, and Rios-Rull (2003), random subjective discount rates of Krusell and Smith (1998). Our paper is also closely related to the work on the importance of entrepreneurship of Quadrini (2000) and Cagetti and De Nardi(2006).

The recent literature has shifted attention to generating a Pareto tailed wealth distribution. Important papers include Benhabib, Bisin, and Zhu (2011, 2015, 2016), Benhabib, Bisin, and Luo (2019), Gabaix et al. (2016), Nirei and Aoki (2016), Cao and Luo (2017), Jones and Kim (2018), Toda (2018), Moll, Rachel, and Restrepo (2019), Sargent, Wang, and Yang (2020), and Hubmer, Krusell, and Smith (2021). This literature applies either the Kesten process (Kesten (1973)) or the random growth process to model the earnings and wealth processes and tries to give microfundations to such processes. We refer readers to Gabaix (2009) and Benhabib and Bisin (2018) for surveys of this literature and additional references.

Unlike this literature, our wealth and income processes belong to the class of AJD processes of Duffie, Pan, and Singleton (2000). Given exponential-type jump size distributions, these processes have exponential tails. We give conditions such that the wealth process has a heavier tail than the labor income process. As exponential-type tails and power-type tails are almost indistinguishable empirically (Heyde and Kou (2004)), our model can generate the wealth distribution on the top close to the data.

[4]See Heathcote, Storesletten, and Violante (2009), Guvenen (2011), Quadrini and Rios-Rull (2015), Krueger, Mitman, and Perri (2015), and Benhabib and Bisin (2018) for recent surveys.

Our modeling of investment risks builds on Angeletos and Calvet (2006), Angeletos (2007), and Angeletos and Panousi (2009). Our model differs from theirs in three major ways in addition to some technical details: (i) We introduce capital return jump risks in addition to labor income risks. (ii) We adopt the recursive utility specification of Weil (1993) in a continuous-time setup, which ensures consumption is always positive. (iii) Our model generates a nondegenerate wealth distribution as in the data, matching the top wealth shares up to the 0.1%.

Our model is also related to Wang (2007), who adopts recursive utility with a consumption-dependent rate of time preference (Uzawa (1968)). Such a specification in an exponential-affine framework allows him to derive a closed-form solution to the individual consumption/saving problem and a moment characterization of the wealth distribution. As discussed earlier, his model suffers from the problem of the standard BHA model and cannot generate realistic income and wealth distributions as in the data.

2 Model

Consider an infinite-horizon continuous-time model in which there is a continuum of infinitely-lived households, indexed by i and distributed uniformly over $[0, 1]$. At each time $t \geq 0$, each household is endowed with one unit of labor. It owns and runs a private firm, which employs labor supplied by other households in the competitive labor market but can only use the capital stock invested by the particular household. Each household faces two independent sources of idiosyncratic shocks that hit its private firm and its earnings. It can only trade riskless bonds and cannot fully diversify away idiosyncratic shocks. We assume that there is no aggregate uncertainty so that all aggregate variables are deterministic by a law of large numbers. We focus on a stationary economy in which all aggregate (per capita) quantities and prices (wage and interest rate) are constant over time.

2.1 Preferences

All households have the same recursive utility over consumption in continuous time (Duffie and Epstein (1992)). It helps intuition much better by motivating such utility as the limit of a discrete-time model (Epstein and Zin (1989)) as the time interval shrinks to zero.

Let dt denote the time increment. The continuation utility U_t at time t over a consumption process $\{c_t\}_{t \geq 0}$ satisfies the following recursive equation:

$$U_t = f^{-1} \left[f\left(c_t\right) dt + \exp\left(-\beta dt\right) f\left(\mathcal{R}_t\left(U_{t+dt}\right)\right) \right], \tag{1}$$

where $\beta > 0$ denotes the rate of time preference, f denotes a strictly increasing time aggregator function, and $\mathcal{R}_t$ denotes a conditional certainty equivalent. Notice that U_t is ordinally equivalent to $f\left(U_t\right)$ for a strictly increasing function f. We adopt the specification of Weil (1993):

$$f(c)=\frac{c^{1-1/\psi}}{1-1/\psi},\;\; \mathcal{R}_t(U_{t+dt})=u^{-1}\mathbb{E}_t u(U_{t+dt}),\; u(U_{t+dt})=\frac{-\exp(-\gamma U_{t+dt})}{\gamma}, \tag{2}$$

where $\gamma>0$ is the coefficient of absolute risk aversion and $\psi>0$ ($\psi\neq 1$) is the EIS parameter. In Appendix B, we derive the continuous-time limit for U_t as $dt\to 0$ in the presence of both jump and diffusion risks. Such a construction ensures dynamic consistency of the continuation utility. By varying the consumption process $\{c_t\}_{t\geq 0}$, we obtain the utility function $U\left(\{c_t\}_{t\geq 0}\right)=U_0$.

The specification of f in (2) implies that consumption can never be negative. Moreover, the CARA specification of u allows the consumption/saving problem with additive labor income risk to admit a close-form solution (Weil (1993)). Angeletos and Calvet (2006) also consider CARA specification for u, but they assume that $f(c)=-\psi\exp(-c/\psi)$ is an exponential function. This specification implies that optimal consumption can be negative and cannot generate a stationary wealth distribution. To ensure the existence of a stationary wealth distribution, Wang (2007) adopts recursive utility with CARA specification for $u=f$ and with consumption-dependent rate of time preference (Uzawa (1968)).[5]

2.2 Decision Problem

In this subsection we study a household's decision problem holding the interest rate and the wage rate constant over time. For simplicity we omit the household-specific index i. Let the production function take the form

$$y_t=Ak_t^{\alpha}l_t^{1-\alpha},\; \alpha\in(0,1),\; A>0,$$

where A, y_t, k_t, and l_t denote total factor productivity (TFP), output, capital, and labor, respectively. Profit maximization implies

$$R^k k_t=\max_{l_t}\{Ak_t^{\alpha}l_t^{1-\alpha}-wl_t-\delta k_t\}=\left[\alpha A\left(\frac{(1-\alpha)A}{w}\right)^{\frac{1-\alpha}{\alpha}}-\delta\right]k_t, \tag{3}$$

where w and R^k denote the wage rate and capital return, and $\delta>0$ denotes the depreciation rate.

The household faces idiosyncratic investment risk and labor income (earning) risk. The effective market hours are represented by ℓ, which is hit by a Brownian shock driven by the dynamics

$$d\ell_t=\rho_\ell(L-\ell_t)\,dt+\sigma_\ell\sqrt{\ell_t}dW_t^{\ell}, \tag{4}$$

where W_t^{ℓ} is a standard Brownian motion and $\sigma_\ell,\rho_\ell>0$. One can interpret ℓ as the product of labor

[5] Another way to generate a stationary wealth distribution is to adopt the overlapping-generations framework of Blanchard (1985) with death probability independent of age, or with "perpetual youth," yielding a decreasing exponential age distribution. As pointed out in Benhabib and Bisin (2018), an implication of the "perpetual youth" assumption is that the right tail becomes populated with agents that are unrealistically old for calibrations that match the right tail of the wealth distribution.

hours and idiosyncratic labor productivity. To ensure ℓ_t is positive, we assume that $2\rho_\ell L \geq \sigma_\ell^2$. The estimation from PSID data later supports this restriction. The long-run mean of ℓ_t is equal to aggregate labor supply L.

The capital return is hit by a jump shock dJ_t, where J_t is a Poisson jump process. For each realized jump, the jump size q is drawn from a fixed probability distribution ν over $[0, \infty)$. Assume that all shocks are independent of each other and across households.

Suppose that the intensity at which a jump occurs depends upon k_t and is given by $\lambda_t = \lambda_k k_t$, where $\lambda_k > 0$. Intuitively, during any time interval $[t, t+dt]$, the household receives an average capital income $\lambda_k k_t \mathbb{E}_\nu (q)\, dt$. The interpretation is that there is a rare event that the investment earns large returns and the success probability is positively related to the capital stock. It represents additional output from entrepreneurial risk taking activities like innovations or R&D.[6] It is related to the early seminal work of Quadrini (2000) which introduces entrepreneurship through stochastically arising profitable investment opportunities for households in a "non-corporate sector' subject to borrowing constraints, and to the work of Cagetti and De Nardi (2006) that also incorporates entrepreneurial entry, exit, and investment decisions in the presence of borrowing constraints in an OLG framework of perpetual youth as in Blanchard (1985). Our modeling is also similar to the "awesome state" or rare event in which individual labor productivity can become extremely high (Castaneda, Diaz-Gimenez, and Rios-Rull (2003)). The key difference is that in our model the large income comes from capital.

Capital assets are illiquid and trading k_t of them incurs adjustment costs given by $\eta k_t^2/2 + \chi k_t$ per unit of time, where $\eta > 0$ and $\chi > 0$ are parameters. The household can also trade riskless bonds at the interest rate r to insure against idiosyncratic shocks. Let b_t denote the household's holding of bonds. Households can borrow and lend among themselves without any trading frictions so that $b_t < 0$ represents borrowing. To deliver a close-form solution, we do not impose binding borrowing constraints, but a transversality condition on the value function must be satisfied to rule out Ponzi schemes (e.g., Merton (1971)).

Let $x_t = b_t + k_t$ denote the household's wealth level. Then the household faces the following dynamic budget constraints

$$dx_t = r x_t dt + \left(R^k - \chi - r\right) k_t dt - \frac{\eta}{2} k_t^2 dt + dJ_t + w\ell_t dt - c_t dt. \tag{5}$$

The household problem is to choose consumption and capital investment processes $(c_t, k_t)_{t\geq 0}$ to maximize utility $U\left(\{c_t\}_{t\geq 0}\right)$ subject to the budget constraints (5), given initial wealth $x_0 = x$ and initial labor $\ell_0 = \ell$.

Let $V(x, \ell)$ denote the value function. In Appendix A, we use dynamic programming to derive the following result:

Proposition 1. *Suppose that* $0 < r < \beta$ *and that* $\mathbb{E}_\nu \exp(-\alpha q)$ *and* $\mathbb{E}_\nu (q)$ *are finite for any* $\alpha > 0$. *Then*

[6] One may assume that the intensity contains a constant component in that $\lambda_t = \lambda_0 + \lambda_k k_t$ for $\lambda_0 > 0$. In this case there is always a chance that the household receives additional output without making any investment.

the optimal consumption rule is given by

$$c_t = \theta^{1-\psi}\left(x_t + \xi_\ell \ell_t + \xi_0\right), \tag{6}$$

the capital demand is given by

$$k_t = k \equiv \frac{1}{\eta}\left[R^k - \chi + \frac{\lambda_k \mathbb{E}_\nu\left(1-\exp\left(-\gamma\theta q\right)\right)}{\gamma\theta} - r\right], \tag{7}$$

and the value function takes the form

$$V\left(x_t, \ell_t\right) = \theta\left(x_t + \xi_\ell \ell_t + \xi_0\right), \tag{8}$$

where θ, ξ_ℓ, and ξ_0 are given by

$$\theta = \left[\psi\left(\beta - r\right) + r\right]^{\frac{1}{1-\psi}}, \tag{9}$$

$$\xi_\ell = \frac{-\left(\rho_\ell + r\right) + \sqrt{\left(\rho_\ell + r\right)^2 + 2\sigma_\ell^2\theta\gamma w}}{\theta\gamma\sigma_\ell^2} > 0, \tag{10}$$

$$\xi_0 = \frac{1}{r}\left\{\left(R^k - \chi - r\right)k - \frac{\eta}{2}k^2 + \frac{\lambda_k k}{\gamma\theta}\mathbb{E}_\nu\left[1 - \exp\left(-\gamma\theta q\right)\right] + \xi_\ell\rho_\ell L\right\}. \tag{11}$$

Equation (7) describes the optimal capital rule similar to that in the portfolio choice literature (Merton (1971)) and can be rewritten as

$$R^k + \lambda_k \mathbb{E}_\nu\left[q\right] - r = \left(\eta k + \chi\right) + \lambda_k\left[\mathbb{E}_\nu\left[q\right] - \frac{\mathbb{E}_\nu\left(1-\exp\left(-\gamma\theta q\right)\right)}{\gamma\theta}\right]. \tag{12}$$

The left side of this equation represents the expected return on capital investment in excess of the interest rate (i.e., equity premium). The expected return consists of the usual return R^k from neoclassical production and the return $\lambda_k \mathbb{E}_\nu\left[q\right]$ from business risk-taking activities. The right side has two components. The first component $\eta k + \chi$ represents the marginal capital adjustment or maintenance cost, which reflects liquidity premium. The second component represents the risk premium due to the jump risk and increases with the risk aversion parameter γ and the jump intensity λ_k. Notice that optimal capital demand k is constant and independent of individual variables and hence will be equal to the aggregate capital stock in equilibrium.

To understand the consumption rule in (6), we need to introduce the concept of human wealth, which is defined as the expected present value of future labor income. For our incomplete markets model with uninsured risk, there is no unique stochastic discount factor used to discount future labor income. The literature typically uses the interest rate $r > 0$ as the discount rate. Formally, we define human wealth as

$$h_t \equiv \mathbb{E}_t\left[\int_t^\infty e^{-r(s-t)} w \ell_s ds\right] = \frac{w}{r + \rho_\ell}\left(\ell_t + \frac{\rho_\ell L}{r}\right). \tag{13}$$

Then we can rewrite (6) as

$$c_t = \vartheta \left(x_t + a_h h_t + \Gamma \right), \tag{14}$$

where we define

$$\vartheta \equiv \psi \left(\beta - r \right) + r > 0 \tag{15}$$

$$a_h \equiv \frac{\left(r + \rho_\ell \right) \xi_\ell}{w} \in \left(0, 1 \right), \tag{16}$$

$$\Gamma \equiv \frac{\eta k^2}{2r}. \tag{17}$$

The variable ϑ represents the marginal propensity to consume (MPC), which is important to understand the consumption behavior and the wealth distribution. The assumption of $0 < r < \beta$ ensures that the MPC $\vartheta > 0$ by equation (15), which also shows that the MPC increases with the EIS parameter ψ. This assumption will be satisfied in general equilibrium. As is well known, the MPC is equal to r in the standard time-additive CARA utility model (e.g., Caballero (1990) and Wang (2007)). Importantly, recursive utility in our model helps generate a MPC higher than r.

It follows from (10) that $a_h \in (0, 1)$. As pointed out by Wang (2007), the square-root process in (4) implies that a higher level of current labor income generates a more volatile stream of future labor incomes. Hence the household's precautionary saving is larger when its labor income level is higher, causing the household to consume less out of its human wealth than out of its financial wealth. Such precautionary saving is given by $\left(1 - a_h \right) h_t$, which is stochastic. The remaining term in (14), $\Gamma = \eta k^2 / \left(2r \right)$, can be rewritten as

$$\Gamma = \frac{1}{r} \left\{ \left(R^k - \chi - r \right) k + \frac{\lambda_k k}{\gamma \theta} \mathbb{E}_\nu \left[1 - \exp \left(-\gamma \theta q \right) \right] - \frac{\eta}{2} k^2 \right\},$$

according to (7). Thus Γ is essentially equal to the present value of expected (risk- and cost- adjusted) profits from the capital investment. The risk adjustment captures precautionary savings against capital jump risks.

2.3 Stationary Equilibrium

We now add household-specific index i, conduct aggregation, and define equilibrium in the steady state. Aggregate consumption, labor, capital, wealth, and output are given by

$$\begin{aligned} C_t &= \int c_t^i di, \; L = \int \ell_t^i di, \;\; K_t = \int k_t^i di, \; X_t = \int x_t^i di, \\ Y_t &= \int y_t^i di + \lambda_k \mathbb{E}_\nu [q] \int k_t^i di. \end{aligned}$$

Aggregate output Y_t consists of two components: total output generated by firm production $\int y_t^i di$ and extra output generated by business risk taking.

A stationary competitive equilibrium consists of constant wage w and interest rate r, individual choices $\{c_t^i, k_t^i, l_t^i\}_{t\geq 0}$ for $i \in [0,1]$, and constant aggregate quantities C, Y, and K, such that (i) given (w, r), the processes $\{c_t^i, k_t^i, l_t^i\}_{t\geq 0}$ are optimal choices for each household i; (ii) the bond and labor markets clear

$$\int b_t^i di = 0, \quad \int l_t^i di = L.$$

According to the constant-returns-to-scale technology in (3), we can show that the capital/labor ratio is identical for all households. Thus we have

$$R^k = \alpha A K^{\alpha-1} L^{1-\alpha} - \delta, \tag{18}$$

$$w = (1-\alpha) A K^{\alpha} L^{-\alpha}, \tag{19}$$

and $AK^{\alpha}L^{1-\alpha} = \int y_t^i di$. Moreover, it follows from (7) that $k = K$. Then we have

$$R^k K + wL = AK^{\alpha}L^{1-\alpha} - \delta K. \tag{20}$$

In the steady state, aggregate wealth satisfies $X = K$. Aggregating the budget constraint (5) and using (20), we can derive the resource constraint

$$C + \delta K + \frac{\eta}{2}K^2 + \chi K = Y, \tag{21}$$

where aggregate output is given by

$$Y = AK^{\alpha}L^{1-\alpha} + \lambda_k \mathbb{E}_{\nu}[q] K. \tag{22}$$

By (14), aggregate consumption is given by

$$C = \vartheta (K + a_h H + \Gamma), \tag{23}$$

where we can write aggregate Γ and aggregate human wealth (from (13)) as

$$\Gamma = \frac{\eta K^2}{2r}, \tag{24}$$

$$H \equiv \int h_t^i di = \frac{wL}{r}. \tag{25}$$

3 Equilibrium Analysis

In this section we first analyze the properties of the stationary wealth distribution taking prices (interest rate and wage) as given and then study the determination of equilibrium prices.

3.1 Wealth Distribution

To study the wealth distribution, we substitute the optimal consumption rule (6) into the wealth dynamics (5) to derive

$$dx_t = -\rho_x x_t dt + \mu_x dt + \phi w \ell_t dt + dJ_t, \tag{26}$$

where

$$\mu_x \equiv \left(R^k - \chi - r\right) k - \frac{\eta}{2} k^2 - \vartheta \xi_0, \tag{27}$$

$$\rho_x \equiv \psi\left(\beta - r\right), \;\; \phi \equiv 1 - \frac{\vartheta \xi_\ell}{w}. \tag{28}$$

Clearly $\rho_x < 0$ if $r < \beta$. The term ϕ represents the marginal propensity to save (MPS) out of labor income. We restrict parameter values such that $\phi > 0$ in equilibrium.

Let $z_t \equiv w\ell_t$ denote labor income. It follows from (4) that

$$dz_t = \rho_\ell \left(Z - z_t\right) dt + \sigma_z \sqrt{z_t} dW_t^l, \tag{29}$$

where $Z \equiv wL$ and $\sigma_z \equiv \sqrt{w}\sigma_\ell$. For $\rho_x > 0$ and $\rho_\ell > 0$, the joint wealth and labor income process $\{x_t, z_t\}$ has a limiting stationary distribution if $\mathbb{E}_\nu \ln\left(1+q\right) < \infty$ (Jin, Kremer, and Rüdiger (2020)). The assumption on the jump distribution ν means intuitively that large jumps are not strong enough to push the process eventually to infinity. By a law of large numbers, the stationary distribution of the joint process gives the cross-sectional stationary distribution of wealth and earnings. This distribution can be derived numerically using the transform analysis of Duffie, Pan, Singleton (2000). Instead of solving for this distribution due to its technical complexity, we first establish an important property of optimal consumption and then provide a moment characterization of the wealth distribution.

Unlike the usual exponential-affine model (e.g, Caballero (1990) and Wang (2007)), our model setup can generate a positive equilibrium consumption process.

Proposition 2. *Suppose that the assumptions in Proposition 1 hold and* $\mathbb{E}_\nu \ln\left(1+q\right) < \infty$*. Then wealth* x_t *has a stationary distribution with the support* $(\mu_x/\rho_x, \infty)$*. If* $\xi_0 + \mu_x/\rho_x > 0$*, then optimal consumption* c_t *in a stationary equilibrium is positive for all* t*.*

Notice that the drift of the financial wealth process x_t is μ_x, which may be negative or positive depending on particular parameter values. As the support of the stationary wealth distribution is $(\mu_x/\rho_x, \infty)$, some poor households can be in debt with negative wealth in the long run if $\mu_x < 0$. By the optimal consumption rule (6) or (14), each household consumes a fraction of its financial and human wealth as well as investment profits at each time. Due to the square-root process specification, labor income is positive. Due to innovations or R&D, the household's financial wealth may jump up randomly. Thus optimal consumption can never be negative in a stationary equilibrium if $\xi_0 + \mu_x/\rho_x > 0$. This assumption ensures that investment profits are large enough to offset debt. By our calibration in Section 4, this condition is indeed satisfied.

As is well known, the stationary distribution of the labor income process (square-root process) is a Gamma distribution. It is positively skewed and has a positive excess kurtosis (leptokurtic) if $\sigma_z^2 > \rho_\ell Z$ or $\sigma_\ell^2 > \rho_\ell L$. A leptokurtic distribution also implies the distribution has a tail fatter than the normal distribution. Because there is no closed-form solution for the stationary wealth distribution, we extend Wang's (2007) method to compute moments by incorporating capital return jump risks. We characterize all moments whenever they exist in closed form by a recursive formula. For a quantitative analysis, we will use simulations to characterize the full wealth and earnings distribution in Section 4.

As the mean of x_t is X and the mean of z_t is Z, let $\widetilde{x}_t = x_t - X$ and $\widetilde{z}_t = z_t - Z$. Then it follows from (26) that the demeaned processes satisfy the dynamics

$$d\widetilde{x}_t = \left[-\rho_x \widetilde{x}_t + \phi \widetilde{z}_t - \lambda_k K \mathbb{E}_\nu \left(q\right)\right] dt + dJ_t, \tag{30}$$

$$d\widetilde{z}_t = -\rho_\ell \widetilde{z}_t dt + \sigma_z \sqrt{\widetilde{z}_t + Z} dW_t^l, \tag{31}$$

where J_t is a Poisson process with intensity $\lambda_k K$. Let the cross moment be

$$M_{m,n} = \mathbb{E}\left[\widetilde{x}_t^m \widetilde{z}_t^n\right], \quad m, n = 0, 1, 2, ...$$

Proposition 3. *Let the assumptions in Proposition 1 hold and let* $\mathbb{E}_\nu \ln\left(1+q\right) < \infty$. *Suppose that* $\zeta_j \equiv \mathbb{E}_\nu \left[q^j\right] > 0$ *is finite for* $1 \leq j \leq j^*$ *and* ζ_j *does not exist for* $j = j^* + 1$. *Then the moments of the joint stationary distribution of* (x_t, z_t) *satisfy the recursive equation for* $0 \leq m \leq j^*$ *and* $n \geq 0$:

$$M_{m,n} = \frac{1}{\kappa_{m,n}} \left[\begin{array}{c} P_0\left(n\right) M_{m,n-1} + P_0\left(n\right) Z M_{m,n-2} + P_1\left(m\right) M_{m-1,n+1} \\ + P_2\left(m\right) M_{m-2,n} + \sum_{j=3}^{m} P_j\left(m\right) M_{m-j,n} \end{array} \right], \tag{32}$$

where $M_{0,0} = 1, M_{0,1} = M_{1,0} = 0,$

$$\begin{aligned} \kappa_{m,n} &\equiv \rho_x m + \rho_\ell n, \\ P_0\left(n\right) &\equiv \frac{1}{2}\sigma_z^2 n \left(n-1\right), \quad P_1\left(m\right) \equiv \phi m, \\ P_2\left(m\right) &\equiv \left(\lambda_k K\right) \binom{m}{2} \zeta_2, \\ P_j\left(m\right) &\equiv \lambda_k K \binom{m}{j} \zeta_j, \quad 3 \leq j \leq m. \end{aligned}$$

The moment $M_{m,n}$ *does not exist for any* $m > j^*$.

To apply this proposition we need to initialize the recursion by computing moments for $0 \leq m < 3$ or $0 \leq n < 2$. We provide the details in Appendix A.

Proposition 4. *The variance ratio of wealth to income is given by*

$$\frac{Var_x}{Var_z} = \frac{\phi^2}{\rho_x\left(\rho_x+\rho_\ell\right)} + \frac{\lambda_k K\zeta_2}{2\rho_x Var_z}, \tag{33}$$

where $Var_z = \sigma_z^2 Z/\left(2\rho_\ell\right)$ is the long-run income variance. The correlation between the wealth and labor income processes is given by

$$\frac{\phi}{\rho_x+\rho_\ell}\sqrt{\frac{Var_z}{Var_x}}. \tag{34}$$

As equation (26) shows, the variable ϕ can be interpreted as the MPS out of labor income. The first term on the right side of (33) is the same as that in equation (53) of Wang (2007). A larger value of MPS out of labor income induces a larger variance ratio of wealth to labor income. In the presence of capital assets, the capital return variability contributes to the wealth variance as well. The second term in (33) reflects this contribution which comes from the random capital return jump. Equation (34) shows that if $\phi > 0$, then wealth and labor income are positively correlated.

To understand the income and wealth inequality, we study the skewness and the kurtosis, denoted by $Skew[x]$ and $Kurt[x]$ for any variable x. In Appendix A we derive the following result.

Proposition 5. *Suppose that $\zeta_j \equiv \mathbb{E}_\nu\left[q^j\right]$ is finite for $1 \leq j \leq 4$. Then*

$$Skew\left[x\right] = Skew\left[z\right]\frac{2\sqrt{\rho_x\left(\rho_x+\rho_\ell\right)}}{2\rho_x+\rho_\ell}\left[1+\frac{\left(\lambda_k K\zeta_2\right)\left(\rho_x+\rho_\ell\right)}{2M_{0,2}\phi^2}\right]^{-3/2} + \frac{\lambda_k K\zeta_3}{3\rho_x\left(M_{2,0}\right)^{3/2}}, \tag{35}$$

and

$$\begin{aligned} Kurt\left[x\right] \quad = \quad & Kurt\left[z\right]\frac{\rho_x\left(5\rho_\ell+6\rho_x\right)}{\left(3\rho_x+\rho_\ell\right)\left(2\rho_x+\rho_\ell\right)}\left[1+\frac{\varpi_1\rho_x\left(\rho_x+\rho_\ell\right)}{\phi^2}\right]^{-2} \\ & +3\frac{\phi^2\left(\rho_\ell\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]+3\phi^2\rho_x\right)}{\left(3\rho_x+\rho_\ell\right)\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]^2} - 3 \\ & +\left[\frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}\frac{\sigma_z^2 Z M_{0,2}\varpi_1}{2\left(\rho_x+\rho_\ell\right)}+\varpi_2\right]\frac{1}{M_{2,0}^2}, \end{aligned} \tag{36}$$

where $\varpi_1 > 0$ and $\varpi_2 > 0$ are given in Appendix A.

If capital and bonds are perfect substitutes (i.e., $\lambda_k = \chi = \eta = 0$), this proposition is reduced to equations (62) and (63) in Wang (2007). In this case the wealth skewness and kurtosis are smaller than the labor income skewness and kurtosis. This result is related to Theorem 8 of Stachurski and Toda (2019), which states that the tail thickness of the model output (wealth) cannot exceed that of the input (labor income) in the standard BHA model. Our model departs from such a standard BHA model by separating illiquid capital assets from liquid assets and by introducing capital return jump risk.

Jump risk does not necessarily generates a more positively skewed and fatter tailed wealth distribution. The first term on the right side of equation (35) shows that the capital return jump risk reduces the wealth skewness relative to the earnings skewness. However, the second term raises the wealth skewness when $\zeta_3 > 0$.

Equation (36) shows that the wealth kurtosis consists of three components. The capital return jump risk reduces the first component (the first line of (36)) as ϖ_1. We can also show that the second component (the second line of (36)) is negative. Only the last component (the third line of (36)) can raise the kurtosis of the wealth distribution because $\varpi_1 > 0$ and $\varpi_2 > 0$. In Appendix A, we show that ϖ_2 is positively related to $\zeta_4 > 0$.

In summary, the capital return jump risk may not generate higher skewness and higher kurtosis for the wealth distribution relative to the labor income. The jump size distribution (the third and fourth moments) is critical to determine the skewness and kurtosis of the wealth distribution. As is well known, skewness and kurtosis may not fully capture the tail behavior of a distribution. In Section 5 we will provide an explicit characterization of the tail behavior of the wealth and labor income distributions.

3.2 Aggregate Investment and Savings

We now use the asset/investment demand and supply analysis of Aiyagari (1994) to understand the aggregate equilibrium determination. A tractable feature of our model is that we do not need to know the full wealth distribution to conduct aggregation. In particular only the mean matters for the aggregate. Thus we can compute stationary equilibrium independent of the full wealth distribution, which is determined after the equilibrium interest rate and wage are determined.

We first derive the aggregate investment demand curve. Combining equations (7) and (18) yields

$$\alpha(K/L)^{\alpha-1} - \chi - \delta = \eta K + r - \frac{\lambda_k \mathbb{E}_\nu \left[1 - \exp\left(-\gamma\theta q\right)\right]}{\gamma\theta}. \tag{37}$$

Because θ is a function of r given in (9), we can use the above equation to derive aggregate capital K as a function of the interest rate r, denoted by $K\left(r\right)$. Then we obtain the aggregate investment demand curve $\delta K\left(r\right)$. The lemma below characterizes the properties of $K\left(r\right)$.

Lemma 1. *Let the assumptions in Proposition 1 hold. Then there is a unique solution to equation (37) for any $r > 0$, denoted by $K\left(r\right)$, which decreases with r and satisfies $\lim_{r\to\beta} K\left(r\right) = K\left(\beta\right)$ and $\lim_{r\to 0} K\left(r\right) = K\left(0\right)$.*

Next we derive the aggregate saving curve. Aggregate savings S are given by

$$S \equiv Y - C - \frac{\eta}{2}K^2 - \chi K = AK^\alpha L^{1-\alpha} + \lambda_k \mathbb{E}_\nu\left[q\right] K - \vartheta\left(K + a_h H + \Gamma\right) - \frac{\eta}{2}K^2 - \chi K, \tag{38}$$

where we have substituted the aggregate consumption function in (23) into the above equation. Since aggregate capital K is a function of r, aggregate output Y is also a function of r. We use (18) and (19) to derive R^k and w as functions of r and hence aggregate consumption is also a function of r by (24) and (25). As a result, S is a function of r.

In Appendix A, we show that

$$\begin{aligned} S(r) &= (r+\delta+\chi-\vartheta)K + wL(1-\vartheta a_h/r) + \frac{1}{2}\eta K^2\left(1-\frac{\vartheta}{r}\right) \\ &\quad +\lambda_k K\left(\mathbb{E}_\nu[q] - \frac{\mathbb{E}_\nu[1-\exp(-\gamma\theta q)]}{\gamma\theta}\right). \end{aligned} \tag{39}$$

Aggregate savings consist of four components: The first component $(r+\delta-\chi-\vartheta)K$ represents savings out of capital assets. The second component is precautionary savings against the Brownian labor income risk. The third component represents savings out of capital returns. The last component represents precautionary savings against the capital return jump risk.

By the market-clearing condition, aggregate saving is equal to aggregate investment so that

$$S(r) = \delta K(r), \tag{40}$$

which determines the stationary equilibrium interest rate. The following lemma characterizes the aggregate saving function:

Lemma 2. *Let the assumptions in Proposition 1 hold. Then*

$$\lim_{r\uparrow\beta} S(r) > \delta K(\beta) \text{ and } \lim_{r\downarrow 0} S(r) = -\infty.$$

The limiting behavior of the saving function is very different from that in Aiyagari (1994). In his model with time-additive power utility and borrowing constraints, $S(r)$ approaches infinity as r increases to β and $S(r)$ approaches to the borrowing limit as r decreases to zero. In our model with recursive utility and without borrowing constraints, $S(r)$ tends to negative infinity as r decreases to zero. That is, households want to borrow as much as possible because there is no borrowing constraint. As r increases to β, the MPC ϑ approaches $\beta = r$. Thus the first component of aggregate savings in (39) approaches $\delta K(\beta)$. But all other components of aggregate savings in (39) are nonnegative and precautionary savings are strictly positive. We thus deduce that aggregate savings are higher than aggregate investment as $r \to \beta$.

Combining Lemmas 1 and 2, we immediately obtain the following result:

Proposition 6. *Suppose that $\mathbb{E}_\nu \exp(-\alpha q)$ and $\mathbb{E}_\nu(q)$ are finite for any $\alpha > 0$. Then there exists an equilibrium with $0 < r < \beta$.*

For comparison, we consider two alternative economies for which we modify the investment demand curve (37) and the saving curve (39):

(i) The complete markets economy with fully insured idiosyncratic risks. In this case we have $\sigma_\ell = 0$ and each household receives $\lambda_k\mathbb{E}_\nu[q]$ as capital return. From (10) and (16), we can show that $a_h = 1$

when $\sigma_\ell = 0$. Then the investment demand $\delta K(r)$ can be derived from the following equation:

$$\alpha A K^{\alpha-1} L^{1-\alpha} - \delta + \lambda_k \mathbb{E}_\nu [q] - \chi - r = \eta K.$$

The saving function becomes

$$S^c(r) = (r + \delta + \chi - \vartheta)K + wL(1 - \vartheta/r) + \frac{1}{2}\eta K^2(1 - \vartheta/r).$$

In steady-state equilibrium, we have $S^c(r) = \delta K(r)$, which leads to $r = \beta$.

(ii) The BHA economy with only labor income risks, in which capital and bond assets are perfect substitutes, i.e., $\eta = \chi = \lambda_k = 0$. In this case investment demand $\delta K(r)$ is determined by

$$\alpha A K^{\alpha-1} L^{1-\alpha} = r + \delta,$$

and the saving function is

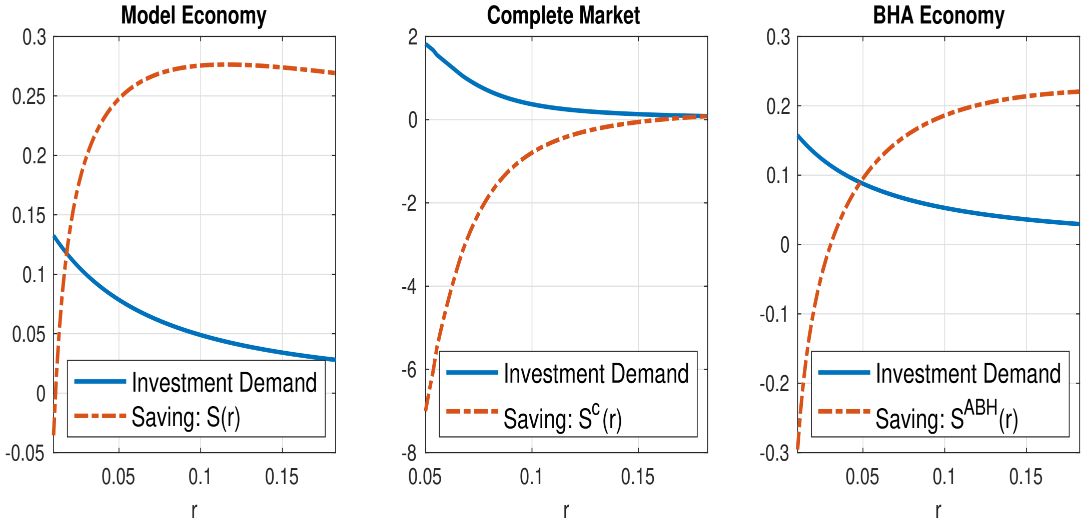
$$S^{BHA}(r) = (r + \delta - \vartheta)K + wL(1 - \vartheta a_h/r).$$

The marginal product of capital is equal to the interest rate plus the depreciate rate. By similar arguments in the proof of Proposition 6, we have $0 < r < \beta$.

Figure 1 illustrates the three economies with the parameters to be discussed in the next section. The investment demand curves are all downward sloping functions of the interest rate r. In the complete markets case, the saving function only intersects with the investment demand curve at the interest rate $r = \beta$. The standard BHA economy features an intersection at $0 < r < \beta$ for precautionary saving reasons. In our model economy, there are additional idiosyncratic investment risks, so the interest rate is even lower. Interestingly, the saving function in our model initially increases but later decreases with the interest rate r. This is due to the different income and substitution effects. The investment curve intersects the saving curve only once, giving one equilibrium solution even when the saving function is hump-shaped.

We close this section by discussing the equilibrium level of capital stock. Angeletos (2007) argues that the investment risk raises precautionary savings, but also decreases investment demand. Thus the net effect on the steady-state capital stock is ambiguous even though $r < \beta$. In our model we have both labor income and investment risks. The steady-state capital stock under incomplete markets is lower than that under complete markets given our calibrated parameter values discussed in the next section. The main reason is that the decrease in investment demand due to large investment (jump) risks dominates the precautionary saving effect.

Figure 1: Investment and Savings functions in three economies



4 Quantitative Results

In this section we calibrate our model and examine its quantitative implications for the aggregate economy and for the income and wealth distributions. We solve for the stationary equilibrium numerically and suppose that one unit of time in our model corresponds to one year.

4.1 Calibration

We group all model parameters in three sets and choose parameter values such that the stationary equilibrium of the calibrated model matches the US macro- and micro-level data.

Standard Parameters. First, consider $\{\alpha, \delta, \psi, \chi, \eta, \beta, \gamma\}$. We set the capital share $\alpha = 0.33$ as in the macro literature. Set the depreciation rate $\delta = 7.33\%$ to target 16.4% investment to output ratio in the US data. We set the EIS parameter $\psi = 1.41$ in line with the finance literature on long-run risk, and later we conduct a sensitivity analysis with respect to ψ. Set the linear adjustment cost parameter χ, the quadratic adjustment cost parameter η, the subjective discount rate β, and the CARA parameter γ to target the following equilibrium variables: the interest rate $r = 1.8\%$, the capital return premium $R^k - r = 1\%$,[7] the MPC $\vartheta = 0.25$ by (15), in line with most of OECD aggregate MPC measures (Carroll, Slacalek, and Tokuoka (2014)), and the coefficient of relative risk aversion $\gamma C = 5$, where C is the aggregate consumption level in the stationary equilibrium.

[7] One interpretation of $R^k - r$ is liquidity premium. It may be measured by the average of the spread between AAA corporate bonds and treasuries of similar maturity after 1984, which is roughly 1% (see Krishnurmuthy and Vissing Jorgensen (2012), Del Negro et al. (2017), and Cui and Radde (2020)). The private equity premium is around 2% according to Angeletos (2007) for compensating idiosyncratic risks.

Table 1: Calibrated Earning Process

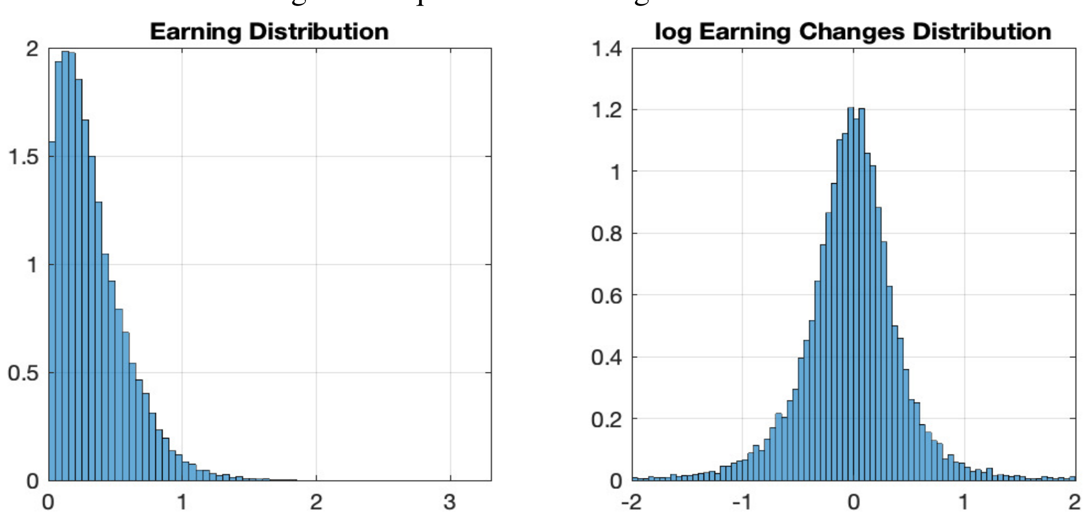
Top	1%	5%	10%	50%	90%
Data	11.0%	22.3%	32.3%	81.5%	99.2%
Model	5.3%	19.0%	31.7%	83.3%	99.3%

Note: statistics are averages between 2000 and 2014 in PSID.

Earnings Process. Second, we calibrate the parameters L, ρ_ℓ, and σ_ℓ in (4) to match some statistics for earnings in the PSID data. We normalize labor supply L to $1/3$ labor hours. We also normalize the steady-state wage rate to one by adjusting the TFP parameter A. As a result, the labor income process (29) is the same as the employment shock process (4). We can then choose ρ_ℓ and σ_ℓ to minimize the distance between the statistics from the process (4) and from earnings in the PSID data. In particular, we target the earning shares for the top 1%, 5%, 10%, 50%, and 90% earners in the PSID data (see Table 1).

Once L is fixed, the square-root process (4) has a known limiting distribution, which is a Gamma type with a shape parameter $2\rho_\ell L/\sigma_\ell^2$ and a scale parameter $2\rho_\ell/\sigma_\ell^2$. Since the Gamma distribution is affected by the ratio ρ_ℓ/σ_ℓ^2 only, there are multiple pairs of ρ_ℓ and σ_ℓ that can generate the same statistics as in Table 1. Thus we also target the variance of annual log earnings changes, 0.23, as in Guvenen et al. (2021) to pin down σ_ℓ. Since the data is annual income, we aggregate the high-frequency individuals' income in computation. Our estimation/calibration procedure is similar to the simulated methods of moments, e.g., Kaplan, Moll, and Violante (2018). The main difference is that we do not need to simulate the labor income process (4) to match the percentiles in the data (Table 1) because the stationary distribution of this process is known. Figure 2 shows the simulated earnings distribution and the distribution for annual changes of log earnings under our calibration. The latter distribution is centred around zero and is almost symmetric, which is consistent with the finding of Guvenen et al. (2021) using the US census data.

Figure 2: Square-root Earning Distribution



It merits emphasis that the square-root earnings process does not feature a fat tail, but it matches reasonably well many statistics in the world-wide inequality database (WWID). In fact, the very top earners have smaller labor income than the others compared to WWID. For example, the average top 0.01% earnings are about 68 times of those of the top 60% people, while the WWID has a factor about 150-200. The average top 5% earnings are 5.2 times of the median earnings, and it is about 5 in the WWID database. Nevertheless, as will be shown later, we can still generate an equilibrium wealth distribution that is much more right skewed and has a much fatter tail than the earnings distribution, due to the capital jump return.

Jump Intensity and Jump Size Distribution. Finally, we consider the remaining parameters in Table 2 that govern the jump process. The jump intensity parameter λ_k is set to 5%, and given the equilibrium capital stock K, the annual probability $\lambda_k K$ of an innovation or R&D is 8%. It should be noted that our model allows both success and failure of innovations or R&D, because the jump returns may not be enough to compensate the loss arising from adjustment costs. This can happen if the jump size is close to zero. We acknowledge that the success probability varies across different sectors and industries. For example in the pharmaceutical industry, the success probability ranges from 4% to 15% across different development stages. Therefore, we experiment with different success probability, and recalibrate parameters. We find that the model implication for the wealth distribution statistics does not change significantly.

The jump size distribution ν is important to match the wealth distribution in the data. Following Cai and Kou (2011), we adopt a hyper-exponential distribution (HED), which is a weighted average of n exponential distributions with nonnegative weights. This type of distributions is flexible and can approximate any completely monotone distributions (Feldmann and Whitt (1998)).[8] The PDF for the HED can be written as

$$f(q) = \sum_{j=1}^{n} p_j \frac{\exp\left(-q/\mu_j\right)}{\mu_j}, \ q > 0, \tag{41}$$

where $p_j \in [0,1]$, $\mu_j > 0$, and $\sum_{j=1}^{n} p_j = 1$. An interpretation is that given an arrival of innovation, a fraction of p_j households draw capital returns from the exponential distribution with mean μ_j.

An important advantage of the HED is that it is analytically tractable. Its moment generating function has a closed form. For the PDF in (41), we have

$$\mathbb{E}_{\nu} \exp(tq) = \sum_{j=1}^{n} \frac{p_j}{1 - \mu_j t}, \tag{42}$$

for $t < \min_j\{1/\mu_j\}$. Thus the Laplace transform $\mathbb{E}_{\nu} \exp(-\alpha q)$ exists and has a simple analytical ex-

[8] Cai and Kou (2011) also study more general mixed-exponential distribution (MED) with possibly negative weights. The MED can approximate any distribution arbitrarily closely (Botta and Harris (1986)). Cai and Kou (2011) show that HED or MED for the jump size is useful for computing option prices given fat-tailed stock returns.

pression for all $\alpha > 0$. This expression will be used in the household decision rules such as (7) and (11). Also notice that the HED in (41) has moments of all orders, which admit a closed form,

$$\mathbb{E}_{\nu}\left[q^{m}\right] = n! \sum_{j=1}^{n} p_j \mu_j^m \quad \text{for } m \geq 1.$$

It follows from Proposition 3 that all moments for the wealth process also exist. Clearly, the wealth distribution in our model does not have a power-law tail. Nevertheless, we show below that our calibrated model can still match reasonably well the wealth shares in the data. In Section 5, we will provide more discussions on this issue.

We consider $n = 3$ components in the HED and choose values of μ_1, μ_2, μ_3, p_2, and p_3 to target five statistics: 14% of the average private returns to innovations and/or R&D (i.e., $\lambda_k \mathbb{E}_{\nu}[q]$ in the model), and top 0.1%, 1%, 10%, and 20% wealth shares in the US data. Griffith (2000) estimates the private returns ranging from 14% to 20% in the US. The public return can be even higher. Our target of 14% for the private return is conservative.

Using administrative tax data, Smith, Zidar, and Zwick (2021) estimate that the top 0.1% and 1% wealth shares increased from 9.9% and 23.9% in 1989 to 15% and 31.5% in 2016, respectively. They also show that the most recent estimates from several approaches in the literature tell starkly different stories about the level and evolution of these wealth shares. We choose 15% and 31.5% in 2016 as our target for the top 0.1% and 1% wealth shares, respectively. According to the average between 2000 and 2019 obtained from the distributional financial account of Federal Reserve Board, the top 1% and 10% wealth shares are 30.2% and 66.7%, respectively. According to the average from the Survey of Consumer Finance after 2000, the top 20% wealth share is 79.8%. We choose 66.7% and 79.8% as our target for the top 10% and 20% wealth shares, respectively.

To compute the wealth shares in our model, we run 100 simulations and compute the average. For each simulation, we discretize the equilibrium wealth process x_t (26). The time increment represents one day. In the end, we run 100 simulations of the wealth process x_t, each simulation having 15 years and 365 days per year and 100,000 people. Increasing simulation length and/or the number of people does not change our results significantly. In (26), important parameters that govern the wealth distributions are $\mu_x = -0.1483$, $\rho_x = 0.2320$, and $\phi = 0.8968$ according to our calibration.

We emphasize that our specification of the HED for the jump size plays an important role due to its flexibility. There is a tension between matching the macro statistics and matching micro statistics, if we adopt a distribution with a limited number of parameters such as the exponential distribution. Given a particular mean in the aggregate, a mixture of distributions gives the model additional parameters to match the cross-sectional distribution statistics without affecting aggregate quantities significantly.

Our choice of the mixture of exponential distributions is parsimonious and tractable. It allows us to match the data reasonably well, especially the top 0.1% wealth share. Of course, the larger the number n of component distributions, the more statistics the model can match. However, as n increases, the

Table 2: Calibrated Parameter Values

	Value	Explanation/Target		Value	Explanation/Target
β	0.1825	MPC = 0.25	μ_1	0.21	top 20% wealth share
γ	8.60	relative risk aversion 5	μ_2	16.4	top 10% wealth share
ψ	1.41	exogenous	μ_3	251.7	top 1% wealth share
α	0.33	capital share	p_1	0.8982	$p_1 + p_2 + p_3 = 1$
δ	0.0733	$I/Y = 0.164$	p_2	0.0978	average innovation return 14%
A	0.8857	$w = 1$	p_3	0.0040	top 0.1% wealth share
η	0.001	$R^k - r = 1\%$	χ	0.0086	interest rate $r = 1.8\%$
L	0.33	normalization	ρ_ℓ	0.0697	estimated from PSID
λ_k	0.05	innovation probability	σ_ℓ	0.2115	estimated from PSID

probability p_j of one of the jumps in the mixed distribution decreases, making the draw from this jump distribution less likely to happen. Therefore, increasing $n = 3$ to $n = 4$ requires substantial more number of people in our simulation which quickly becomes infeasible for standard computing facilities, while the model performance is not substantially improved.[9]

4.2 Results and Sensitivity Analysis

Table 3 presents the baseline quantitative results based on our calibration. Though not targeted, our model generates about 0.5% wealth share for the bottom 50%, slightly smaller than the data 1.7%. Our model also generates 33.2% wealth share for the top 50% to 10%, slightly larger than the data 32%. Given that agents in our model do not face borrowing constraints and there is no tax or transfer in the benchmark simulation, this result is acceptable.

As is well known in the literature, it is notoriously challenging to match the extreme right tail of the wealth distribution in the data. The baseline simulation of our model generates 15.0% wealth share for the top 0.1% people. Compared to the previous literature, our result is much closer to the data. For example, the model of Kaplan, Moll, and Violante (2018) generates 2.3% of liquid wealth and 7% of illiquid wealth held by the top 0.1% people. Even with a Pareto-tailed wealth distribution, the model of Cao and Luo (2017) generates a top 0.1% wealth share of 11%. Notice that our model-generated top 1% wealth share is higher than the data but still not too far. The reason is that our adopted HED specification of the jump size distribution implies that the wealth shares of the top 0.1% and 1% are closely related. We choose to let the model hit the top 0.1% wealth share, which is considerably harder, at the cost of allowing the top 1% wealth share to be slightly higher than the data.

Our model-generated wealth Gini coefficient is 0.77, slightly smaller than the data (0.80), again because some households are in debt in our simulations and these households do not receive government

[9] The simulation has been tested in Mac Book Pro 2020 i7 with 16GB memory. The simulation of the labor income process is around 90 seconds and the simulation of the wealth distribution is around 15 seconds.

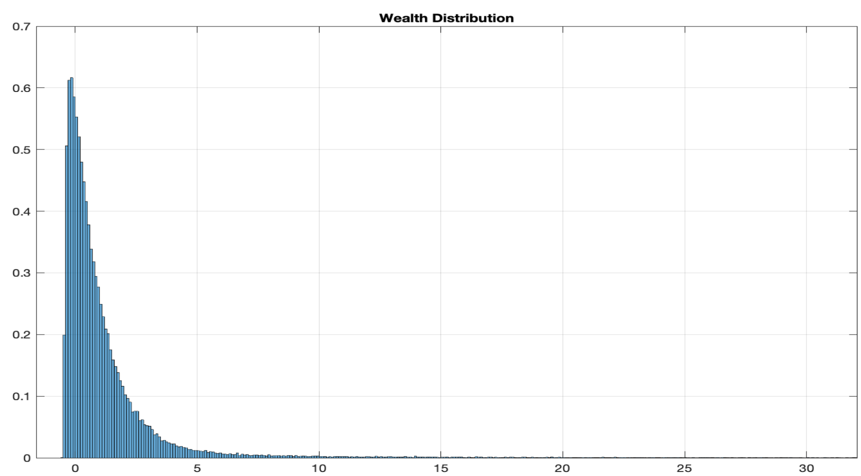
Table 3: Wealth Distribution Statistics

Wealth	Top 0.1%	Top 1%	Top 10%	Top 20%	Top 50-10%	Bottom 50%	Gini
US Net worth	15%	31.5%	66.7%	79.8%	32.0%	1.7%	0.80
Model	15.0%	33.4%	66.0%	80.1%	33.2%	0.5%	0.77

Note: The measure of top 0.1% share is from Smith et.al. (2020). Top 20% data is the average from survey of consumer finance after 2000. The rest are the averages between 2000 and 2019 obtained from distributional financial account of Federal Reserve Board. The model statistics is the average of 100 simulation of 15 years with dt approximated by one day, and each simulation has 100K people starting from the same initial level of labor income and wealth.

transfer.[10] As can be seen from the wealth distribution in Figure 3, there are roughly 25% people having negative wealth, although none of them has negative consumption by Proposition 2.

Figure 3: Wealth Distribution Histogram



How do key parameters affect the result presented in Table 3? To address this question, we conduct a sensitivity analysis by changing values of some key model parameters: the EIS parameter ψ, the risk aversion parameter γ, the Poisson arrival rate λ_k, and the largest mean jump size μ_3 for a component exponential distribution (see Table 4).[11] When changing one parameter value, we keep other parameter values fixed as in the benchmark calibration in Table 2. The sensitivity analysis also helps us better understand our model mechanics.

We first examine the impact of the EIS ψ. Intuitively, consumption/savings behavior depends on the EIS. On the one hand, the higher the EIS ψ, the larger the substitution effect and thus the lower the

[10]There is a well-known issue of calculating the Gini coefficient for a distribution with negative realizations. In our case, the negative wealths from some borrowing agents are all treated as zero. If we move the whole distribution to the right by a constant such that the lowest wealth is zero (instead of a negative number), the Gini will be 0.81.

[11]We also experiment with the probability of the 3rd jump p_3 conditioning on arrival. However, since there is a restriction of $p_1 + p_2 + p_3 = 1$, either p_1 or p_2 needs to vary at the same time. Nonetheless, increasing p_3 has a similar effect both on the aggregate and on the distribution to that of increasing μ_3

savings under incomplete markets. As a result, the equilibrium interest rate r rises with the EIS ψ, ceteris paribus. This can be seen from the downward shift of the saving curve in the left panel of Figure 1. On the other hand, the capital/investment demand changes with ψ by (7) or (12) as $\theta = [\psi(\beta - r) + r]^{\frac{1}{1-\psi}}$ in (9) depends on ψ. The value of θ is very sensitive to a small change of ψ under our calibration compared to varying other parameters. We find that θ declines as ψ increases and hence the jump risk premium in (12) falls with ψ. But the aggregate demand for capital slightly increases with ψ under our parameterization by (37). It turns out that the fall in savings dominates so that the equilibrium interest rate r and the capital return R^k increase with ψ, but the equilibrium capital stock declines with ψ. Aggregate output and consumption also decline with ψ.

We can also look at the MPC $\vartheta = \psi(\beta - r) + r$ in our model: an increase in ψ raises the MPC for a fixed $r \in (0, \beta)$.[12] Intuitively, an increase in the MPC reduces saving incentives even though capital demand increases. Notice that there is a general equilibrium price feedback effect as a higher r reduces ϑ given $\psi > 1$. In our numerical experiment, the former direct effect dominates. Therefore, households borrow more with a higher ψ, and the general equilibrium must feature a higher interest rate to clear the bond market.

The impact of an increase in ψ on wealth inequality depends on four channels. First, a reduction in saving because of a high MPC reduces wealth inequality; this also reduces capital stock, causing innovations to be less likely, and hence reducing capital jump returns and wealth inequality. Second, a decreased capital stock raises the marginal product of capital R^k and hence raises wealth inequality. Third, an increased equilibrium interest rate r raises wealth of the rich and middle class households, but reduces wealth of the poor borrowers, commonly known as the income/wealth effect. At the same time, households will save more or reduce borrowing, in facing a higher interest rate, commonly known as the substitution effect. Fourth, a decreased capital stock reduces the marginal product of labor or the equilibrium wage rate w and thus raises wealth inequality. The net effect is ambiguous. Table 4 shows that the top 10%, 1%, and 0.1% wealth shares as well as the wealth Gini coefficient all rise with ψ in general equilibrium. Therefore, the net effect is dominated by the contribution of a rising interest rate r and falling wage rate w as capital stock declines with ψ.

When we fix the interest rate r and the wage rate w in partial equilibrium, the capital return R^k in (3) as a function of w is also fixed. By (7) or (12), capital demand increases with ψ as the jump risk premium decreases with ψ. We find that wealth inequality increases even more with ψ as the last three channels vanish. Working in the opposite direction of the first channel generates increased wealth inequality due to increased aggregate capital. Moreover, the bottom 50% people are in debt; for example, their debt is 4.5% of aggregate wealth for $\psi = 1.45$. This is in contrast to the case in general equilibrium, in which the increased interest rate as ψ increases induces the poor people to save so that their wealth share becomes 0.4%.

Next we consider the impact of the risk aversion parameter γ. Table 4 shows that the impact of a

[12] By assuming $r > \beta$, Weil (1993) shows that the MPC declines as ψ increases.

Table 4: Sensitivity Analysis

	Capital	Wealth	$r(\%)$	MPC(%)	Bottom 50% (%)	Top 10% (%)	Top 1% (%)	Top 0.1% (%)	Gini (%)
Benchmark	1.60	1.60	1.80	25.00	0.5	66.0	33.4	15.0	76.8
ψ =1.45	1.52	1.52	2.19	25.48	0.4	66.5	34.1	15.2	77.1
– and PE	1.69	1.44	1.80	25.66	-4.5	72.2	37.4	16.8	80.5
ψ =1.5	1.42	1.42	2.68	26.04	0.2	67.1	34.5	15.6	77.5
– and PE	1.81	1.26	1.80	26.48	-12.4	82.5	43.2	19.0	84.5
γ =6	1.55	1.55	2.05	24.90	-0.2	66.2	33.6	14.8	77.1
– and PE	1.69	1.51	1.80	25.00	-3.6	70.7	36.1	15.9	79.7
γ =4	1.48	1.48	2.39	24.76	-1.2	67.6	34.6	15.2	77.9
– and PE	1.83	1.39	1.80	25.00	-9.9	79.5	41.3	18.0	83.6
λ_k =0.06	1.68	1.68	1.49	25.13	-3.0	71.6	36.5	15.9	80.2
– and PE	1.64	1.82	1.80	25.00	1.2	66.5	33.5	14.5	77.1
λ_k =0.07	1.74	1.74	1.24	25.23	-6.8	77.5	39.1	16.5	83.1
– and PE	1.68	2.05	1.80	25.00	1.8	67.7	34.0	14.7	77.8
μ_3 =300	1.63	1.63	1.68	25.05	-0.9	68.6	37.0	17.6	78.6
– and PE	1.60	1.67	1.80	25.00	0.5	66.8	35.7	16.8	77.4
μ_3 =350	1.66	1.66	1.57	25.09	-2.4	71.9	40.7	20.3	80.6
– and PE	1.60	1.74	1.80	25.00	0.6	67.8	37.8	18.8	78.1

Note: For each parameter, the corresponding row shows the result in the stationary equilibrium; the row "- and PE" under denotes the corresponding economy with factor prices r and w fixed.

decrease in γ is similar to that of an increase in the EIS ψ. The main difference is that an increase in γ raises the precautionary saving incentives and pushes down the interest rate, so the MPC $\vartheta = \psi(\beta - r) + r$ increases with γ as $\psi > 1$. As a result, the MPC falls when we reduce γ. Unlike in the case for the EIS ψ, the aggregate and distributional effects are much less sensitive to changes in γ.

Lastly we examine the role of jump risks by focusing on the jump intensity λ_k as well as the component of jump size distribution with the largest mean μ_3. Indeed, the jump risks contribute to rising inequality. In partial equilibrium with fixed prices, both a higher λ_k and a higher μ_3 imply more inequality. Households save more according to the first-order condition for capital k. This can been seen from the larger wealth share of the bottom 50%. However, more frequent jumps still imply more wealth concentrated at the very top, leading to more inequality. In general equilibrium, since investment goes up with either a higher λ_k or a higher μ_3, the interest rate falls below the baseline. Therefore, the poor households save less or borrow more, contributing to even more inequality. The MPC slightly rises with a higher r and a fixed EIS $\psi > 1$, but the aggregate saving still goes up to utilize the better jump returns.

4.3 Application: Taxing the (Super) Rich

In this subsection we study an application of our model by introducing one type of taxation: taxing the super rich households who benefit from the potential large jump returns on investment. This type of policies has gained recent interest because of rising inequality over the past decades.[13] We consider two ways of distributing tax revenues. The first is to transfer tax revenues evenly to all households. The second is to use tax revenues to inject credit to the bond market. We examine the aggregate and redistribution effects.

Specifically, let τ_3 be the proportional tax rate levied on the capital jump return drawn from the third component distribution with the largest mean. Thus on expectation a household receives returns $(1-\tau_3)\mu_3$ per unit of capital if the jump size is drawn from the third component exponential distribution. Total tax revenues are $T = \tau_3 \lambda_k K p_3 \mu_3$.

When these tax revenues are transferred to each household evenly, we can show that the household's problem is modified such that ξ_0 in (11) becomes

$$\xi_0 = \frac{1}{r}\left\{\left(R^k - r\right)k - \frac{\eta}{2}k^2 + \frac{\lambda_k k}{\gamma\theta}\mathbb{E}_\nu\left[1 - \exp\left(-\gamma\theta q\right)\right] + \xi_\ell \rho_\ell L + T\right\},$$

to accommodate the transfer T. This policy also affects the term $\mathbb{E}_\nu\left[1 - \exp\left(-\gamma\theta q\right)\right]$ to reflect taxation. For the aggregate consumption function in (23), the component Γ becomes

$$\Gamma \equiv \frac{\eta k_t^2}{2r} + \frac{T}{r}.$$

Table 5 shows the outcomes of a $\tau_3 = 20\%$ tax rate with equal lump-sum transfer to each household. We first consider the solution in partial equilibrium when prices are fixed at the pretax level. We find the aggregate capital stock K declines as expected, but the change is negligible. The reason is that capital demand in (7) is affected by the change of $\mathbb{E}_\nu\left[1 - \exp\left(-\gamma\theta q\right)\right]$, which is very small as our calibrated $\lambda_k = 0.05$ and $p_3 = 0.004$. Intuitively, the very rich would still invest as much whether they earn \$600M or \$100M a year, as long as it keeps them at the top.

Importantly, the government transfer provides self-insurance and reduces precautionary savings (which affects ξ_0 and the wealth accumulation process), leading the poor households to borrow. Because there is no borrowing constraint in our model, the bottom 50% people are in debt of 67.6% of total wealth and the top 1% people own even more wealth with their wealth share rising from 33.4% to 72.4%. Unlike the Aiyagari (1994) model with a fixed borrowing limit, our model implies that the support of the wealth distribution shifts down. Aggregate wealth X also declines from 1.60 to 0.65. The wealth Gini coefficient increases from 0.768 to 0.90. Thus the policy of taxing the super rich and redistributing evenly

[13] See New York Times article "Private Inequity: How a Powerful Industry Conquered the Tax System" by Jesse Drucker and Danny Hakim (September, 2021) and Propublica article "The Secret IRS Files" by Jesse Eisinger, Jeff Ernsthausen, and Paul Kiel (June, 2021).

raises wealth inequality in partial equilibrium.[14]

We find this result is robust to general equilibrium. But the increase in wealth inequality is less severe. The wealth shares of top 1% and top 0.1% decline, since the tax policy is designed to reduce their wealth. But the top 10% wealth share increases, and the bottom 50% people are still in debt. The wealth Gini coefficient increases from 0.768 to 0.783.

The main reason is the indirect price feedback effect. In general equilibrium, the interest rate must rise to reduce borrowing because the net bond supply is zero. As a result, aggregate capital K is equal to aggregate wealth X in equilibrium and both levels decline from the benchmark level 1.60 to 1.34 after the taxation. Thus, the arrival rate $\lambda_k K$ of capital return jumps declines, thereby reducing wealth inequality relative to the case in partial equilibrium. Though the wage rate declines as K declines, this negative effect on wealth inequality is dominated because of the higher debt incurred by the poor.

Table 5: Taxing Capital Return

	Capital	Wealth	$r(\%)$	MPC(%)	Bottom 50% (%)	Top 10% (%)	Top 1% (%)	Top 0.1% (%)	Gini (%)
Benchmark	1.60	1.60	1.80	25.00	0.5	66.0	33.4	15.0	76.8
$\tau_3 = 0.2$, transfer	1.34	1.34	3.15	24.45	-4.1	68.8	32.9	13.2	78.3
– and PE	1.60	0.65	1.80	25.00	-67.6	141.6	72.4	29.6	90.0
$\tau_3 = 0.2$, injection	1.60	1.53	1.80	25.00	0.47	63.9	30.4	12.3	75.6
– and PE	1.60	1.53	1.80	25.00	0.45	64.3	30.7	12.5	75.8

Notes: For each parameter, the corresponding row shows the result in the stationary equilibrium; the row "- and PE" under denotes the corresponding economy with factor prices r and w fixed.

Now we consider the second distribution policy by injecting tax revenues to the credit market. This policy changes the bond market clearing condition from $B = 0$ to $B + T = 0$. Then this policy only changes the term $\mathbb{E}_\nu \left[1 - \exp\left(-\gamma\theta q\right)\right]$ to reflect taxation in the household consumption/investment problem. As in the first policy experiment with small $\lambda_k = 0.05$ and $p_3 = 0.004$, the policy here does not change aggregate capital K much in partial equilibrium. Aggregate wealth satisfies $X = K+B = K-T$ and is equal to 1.53. Unlike the first transfer policy, the second one reduces wealth inequality because the second policy does not affect households precautionary incentives. In other words, now poor households do not receive transfers that provide self-insurance, so they do not increase their borrowings. We find

[14]In a partial equilibrium model without considering distribution of tax revenues, Benhabib et al. (2011) show that capital tax on r and R^k reduces wealth inequality when wealth follows a Kesten process with a stochastic rate of return. The reason is that the thickness of the right tail depends on the probability of the stochastic wealth return (net of consumption out of wealth) that is above 1 in discrete time and above 0 in continuous time. Inequality is the result of some lucky agents that get long streaks of realizations of high returns. Taxing the stochastic return of wealth/capital reduces the range of its realized values and therefore the inequality. In our framework the rate of wealth return is not stochastic. The policy of taxing capital and redistributing tax revenues evenly reduces precautionary savings of the poor, leading to more inequality. This policy also reduces the equilibrium capital stock and the wage rate and thus disproportionately affects those with lower wealth, further increasing inequality. Such a result is not reported here and available upon request.

the above partial equilibrium results barely change in general equilibrium. The main reason is that the equilibrium interest rate does not change much as the second injection policy does not affect aggregate savings and investment much. Moreover, the equilibrium capital stock barely changes and hence the wage rate does not change much. As a result, the equilibrium price feedback effect is muted.

5 Understanding Exponential versus Power Tails

The previous quantitative results highlight that capital return jump risk is important to generate a realistic wealth distribution as in the data, given a realistic square-root earning process. To understand the intuition, it is critical to study the tail behavior of the wealth distribution. In this section we establish an important result that both the earnings and wealth distributions in our model have an exponential (right) tail instead of a power tail and the wealth tail decays more slowly than the earnings tail.

5.1 Definition of Tail Thickness

We first use the moment generating function to define tail thickness of random variables (Stachurski and Toda (2019)).

Definition 1. *Let X be a random variable and $h(\alpha) = \mathbb{E}[\exp(\alpha X)]$ denote its moment generating function. If $h(\alpha) = \infty$ for all $\alpha > 0$, then X is (right) heavy-tailed. If $h(\alpha)$ is finite for some $\alpha = \alpha_0 > 0$, then X is light-tailed.*[15] *If $h(\alpha)$ is finite for all $\alpha \in [0, \alpha_0]$ for some $\alpha_0 > 0$ and $h(\alpha) = \infty$ for all $\alpha > \alpha_0$, then X has an exponential tail. The exponential decay rate of X is defined as*

$$\overline{\alpha} \equiv \sup\{\alpha \geq 0 : h(\alpha) < \infty\}.$$

If there is a positive exponent λ called the tail index such that

$$\lim_{x \to +\infty} \Pr(X > x) \sim x^{-\lambda},\ \lambda > 0,$$

then X is (right) fat-tailed or Pareto-tailed, where $\sim$ means same up to a constant. If X has finite moments of all orders, i.e., the tail index $\lambda = \infty$, then X is thin-tailed.

We define the tail behavior of a random variable synonymously with that of its distribution. It can be shown that the tail probability $\Pr(X > x)$ of a light-tailed random variable X is bounded above by an exponential function (e.g., Lemma 2 of Stachurski and Toda (2019)). Then Stachurski and Toda (2019) show that

$$\limsup_{x \to \infty} \frac{1}{x} \ln \Pr(X > x) = -\overline{\alpha}.$$

[15] It can be shown that $h(\alpha)$ is finite for all $\alpha \in [0, \alpha_0]$.

This equation gives the intuitive meaning of the exponential decay rate $\overline{\alpha}$, which characterizes the decay speed of the tail probability. For light-tailed distributions whose tails decay faster than any exponential distribution, we have $\overline{\alpha} = \infty$.

A heavy-tailed distribution has a tail that decays more slowly than that of any exponential distribution because

$$\lim_{x \to \infty} \exp(\alpha x) \Pr(X > x) = \infty \text{ for all } \alpha > 0.$$

Then its exponential decay rate $\overline{\alpha} = 0$. Fat-tailed distributions are a subclass of heavy-tailed distributions. Their tails follow a power law or Pareto distribution. By Definition 1, the normal distribution is both light- and thin-tailed. The Pareto distribution is both heavy- and fat-tailed. The log-normal distribution is heavy-tailed, but not fat-tailed. It has moments of all orders. A distribution with tail index λ has finite moments of all orders up to the largest integer below λ.

5.2 Pareto Tail

To study the income and wealth distributions using BHA-type incomplete markets models, one often specifies some exogenous state processes that drive the labor income or capital return fluctuations and then individuals make consumption/saving choices. In general equilibrium, competitive markets determine the interest rate, the wage rate, and the wealth distribution. It turns out that such models are notoriously difficult to match the wealth distribution in the data, especially the wealth shares of the very top percentiles.

To understand our model mechanism and its connection to the literature in a simple unified way, we suppose that there is an exogenous scalar state process z_t satisfying

$$dz_t = \mu^z(z_t)\, dt + \sigma^z(z_t)\, dW_t^z + dJ_t^z,$$

where W_t^z is a standard Brownian motion and J_t^z is a jump process.

Let the wealth process follow the dynamics

$$dx_t = R_t^x x_t dt + y_t dt - c_t dt, \tag{43}$$

where R_t^x denotes the rate of wealth return and y_t denotes labor income that is driven by the state z_t. In the standard BHA model, capital assets and bonds are perfect substitutes and both earn a constant return r and thus $R_t^x = r$. For either CRRA or CARA utility, optimal consumption typically takes the following form

$$c_t = \vartheta x_t + \Phi_t,$$

where Φ_t depends on labor income y_t and ϑ denotes the MPC, which may be different from (15) depending on model setup. In this case (43) becomes

$$dx_t = (r - \vartheta) x_t dt + (y_t - \Phi_t) dt. \tag{44}$$

If $r < \beta$ in equilibrium, we typically have $r < \vartheta$ so that x_t has a stationary distribution. Because randomness of x_t comes from labor income y_t only, the right tail of the wealth distribution is determined by $(y_t - \Phi_t)$. Using the results of Grey (1994), Benhabib and Bisin (2018) and Benhabib, Bisin, and Zhu (2017) show that under some standard BHA assumptions, the wealth distribution inherits the tail behavior of the labor income process (e.g., light-tailedness or the Pareto exponent). Stachurski and Toda (2019) also prove a very related result in discrete time.

In order to generate a wealth distribution that is fatter or heavier and more skewed than the labor income distribution and that even has a fat (Pareto) tail, the literature typically adopts the following two approaches.

1. Kesten process. Saporta and Yao (2005) study the following continuous-time counterpart of Kesten (1973),

$$dx_t = R(z_t) x_t dt + \sigma(z_t) dW_t^x,$$

where z_t is a Markov jump process. They show that if $\mathbb{E}[R(z_t)] < 0$ and $\Pr(R(z_t) > 0) > 0$, then x_t has a stationary distribution with a Pareto tail. If $R(z) < 0$ for all discrete states z, then x_t has a stationary distribution that has finite moments of all orders. In discrete time, the Pareto tail is generated by stochastic discrete shocks to the return on wealth such that the wealth return (net of the fraction of wealth consumed) can exceed 1 with positive probability. Benhabib, Bisin, and Zhu (2011, 2015) apply the Kesten process in discrete time to generate a wealth distribution with a Pareto right tail.

2. Random growth process. Gabaix (2009) and Gabaix et al. (2016) study the following process

$$d \ln x_t = \mu dt + \sigma dW_t^x + dJ_t^x,$$

where W_t^x is a standard Brownian motion and J_t^x is a jump process. They discuss microfoundations of the random growth process and several ways to stabilize the process so that a stationary distribution exists. The mathematical logic of generating a Pareto tail is similar to that for the Kesten process. For the random growth process, the Brownian motion or jumps drive the fluctuations of the growth rate (or change in logarithm) of x_t, while the local expected growth rate $R(z_t)$ itself is driven by Markov jumps z_t for the Kesten process.

For both types of processes, we need the wealth return to be random. Since the equilibrium interest rate is constant, it is important to separate illiquid capital assets from liquid bond assets and then one can introduce randomness (jumps or Brownian motion) to capital returns. For tractability, in this paper we adopt the AJD process introduced by Duffie, Pan, and Singleton (200) in the finance literature. We introduce capital return jumps that appear in equation (26) in the form of disturbances to the change of the wealth level, but not to the change of the log level or the growth rate. Wang (2007) also applies AJD processes. Unlike us, he does not separate capital and bond assets so that (44) still holds. By computing moments explicitly, he shows that the equilibrium wealth distribution is counterfactually less skewed and thinner than the income distribution (also see our Proposition 5).

5.3 Exponential Tail

Given the HED (41) for the jump size $q > 0$, we deduce that $\mathbb{E}_\nu \ln(1+q) < \infty$. It follows from Jin, Kremer, and Rüdiger (2020)) that the joint process $\{x_t, z_t\}$ has a stationary distribution and its law converges to this distribution exponentially fast. As is well known, the square-root labor income process z_t has a stationary Gamma distribution, which has an exponential tail. To study the tail property of the stationary distribution for wealth x_t, we analyze the exponential moments of x_t as $t \to \infty$, $\lim_{t\to\infty} \mathbb{E}[\exp(\alpha x_t)]$ for $\alpha > 0$, following Keller-Ressel and Mayerhofer (2014) and Glasserman and Kim (2010). Notice that when a limiting stationary distribution exists, the limiting exponential moment of x_t does not depend on the initial value x_0.

Proposition 7. *Given the HED (41) for the jump size and $2\rho_\ell L \geq \sigma_\ell^2$, both the stationary wealth and labor income distributions have an exponential tail.*

The macroeconomics literature often uses the Pareto (power law) tail distribution to describe the tail behavior of the wealth or income distribution. Heyde and Kou (2004) argue that it may be very difficult to distinguish empirically the exponential-type tails from power-type tails, even for a sample size of 5000 (corresponding to about 20 years of daily stock return data), although it is quite easy to detect the differences between them and the tails of a normal density. Our quantitative results in Section 4.2 show that our wealth process can match the data reasonably well even if it does not have a Pareto tail.

Both wealth and labor income distributions have an exponential tail. The following proposition characterizes the exponential decay rates of the stationary wealth and labor income distributions.

Proposition 8. *Suppose that the jump size follows the HED (41) and $\phi > 0$ in a stationary equilibrium. Let $2\rho_\ell L \geq \sigma_\ell^2$. Then both the stationary wealth and labor income distributions have exponential tails with the exponential decay rates given by*

$$\overline{\alpha}_x \equiv \min\left\{g(0), \min_j \{1/\mu_j\}\right\}, \quad \overline{\alpha}_z \equiv \frac{2\rho_\ell}{\sigma_\ell^2},$$

respectively, where g is a function defined in the proof of this proposition in Appendix A.

Given our baseline calibration in Table 2, we find that $\overline{\alpha}_x = 1/\mu_3 = 0.004$ and $\overline{\alpha}_z = 3.116$. Thus the wealth distribution has a much smaller exponential decay rate than the labor income distribution.

Intuitively, the exponential decay rate of the wealth distribution depends on that of the capital return jump size distribution. If the capital return jump size is drawn from some exponential distribution with a sufficiently large mean μ_j, then the exponential decay rate of the wealth distribution is given by $1/\mu_j$, which can be much smaller than the exponential decay rate of the income distribution $\overline{\alpha}_z$. The larger is μ_j, the smaller is the exponential decay rate of the wealth distribution. Intuitively, the top wealth share is essentially determined by those who receive large capital return jumps. Without investment risks, the wealth distribution would have a lighter tail than the income distribution (Benhabib and Bisin (2018) and Stachurski and Toda (2019)).

6 Conclusion

In this paper we have provided a tractable heterogeneous-agent model with incomplete markets in continuous time. Departing from the literature which relies on the Pareto-tailed wealth distribution to explain wealth inequality in the data, our model generated wealth distribution has an exponential tail. Because Pareto-tailed distributions and exponential-tailed distributions are almost indistinguishable in data with finite sample, our calibrated model can match the wealth distribution in the data. Our key story is that rich people can build wealth from rare capital return/income jumps through technology innovations or R&D and the jump size is stochastic. The jump size distribution is important to explain the wealth distribution in the extreme right tail.

As an application of our model, we study a fiscal policy by taxing the capital income jump of the super rich. When tax revenues are transferred to all households evenly, such a policy raises wealth inequality. But when tax revenues are used to supply credit to borrowers, such a policy reduces wealth inequality.

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, 2020, Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. forthcoming in Review of Economic Studies.

Aiyagari, S. Rao, 1994, Uninsured Idiosyncratic Risk and Aggregate Saving, Quarterly Journal of Economics 109 (3):659-684.

Angeletos, George-Marios, 2007, Uninsured Idiosyncratic Investment Risk and Aggregate Saving, Review of Economic Dynamics,

Angeletos, George-Marios, and Laurent-Emmanuel Calvet, 2006, Idiosyncratic Production Risk, Growth and the Business Cycle, Journal of Monetary Economics 53, 1095-1115

Angeletos, George-Marios, and Vasia Panousi, 2011, Financial Integration, Entrepreneurial Risk and Global Dynamics, Journal of Economic Theory 146, 863-896.

Benhabib, Jess and Alberto Bisin, 2018, Skewed Wealth Distributions: Theory and Empirics. Journal of Economic Literature 56, 1261-91.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu, 2011, The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents, Econometrica 79, 123-157.

– 2015. The Wealth Distribution in Bewley Economies with Capital Income Risk. Journal of Economic Theory 159, 489-515.

– 2016. The distribution of wealth in the Blanchard-Yaari model. Macroeconomic Dynamics 20. 466-481.

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

Bewley, Truman, 1980, The Optimum Quantity of Money, In *Models of Monetary Economies*, edited by John H. Kareken and Neil Wallace. Minneapolis: Federal Reserve Bank.

Blanchard, Olivier J., 1985, Debt, Deficits, and Finite Horizons, Journal of Political Economy 93, 223-247.

Bricker, Jesse, Alice Henriques, Jacob Krimmel, and John Sabelhaus, 2016, Measuring Income and Wealth at the Top Using Administrative and Survey Data, Brookings Papers on Economic Activity 1, 261-331.

Caballero, Ricardo J., 1990, Consumption Puzzles and Precautionary Savings, Journal of Monetary Economics 25, 113–136.

Cagetti, Marco and Mariacristina De Nardi. 2006. Entrepreneurship, Frictions, and Wealth. Journal of Political Economy 114, 835-870.

Cai, Ning, and S. G. Kou, 2011, Option Pricing Under a Mixed-Exponential Jump Diffusion Model, Management Science 57, 2067–2081.

Cao, Dan, and Wenlan Luo, 2017, Persistent Heterogeneous Returns and Top End Wealth Inequality, Review of Economic Dynamics, 26, 301-326.

Carroll, Christopher D., Jiri Slacalek, and Kiichi Tokuoka, 2014, The Distribution of Wealth and the MPC: Implications of New European Data, The American Economic Review, 104(5), 107–111.

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull, 2003, Accounting for the U.S. Earnings and Wealth Inequality. Journal of Political Economy 111, 818-857.

Cox, John C., Jonathan E. Ingersoll and Stehpen A. Ross, 1985, A Theory of the Term Structure of Interest Rates, Econometrica 53, 385–407.

Cui, Wei, and Sören Radde, 2020, Search-based Endogenous Asset Liquidity and the Macroeconomy, Journal of the European Economic Association 18, 2221-2269.

De Nardi, Mariacristina, 2004, Wealth Inequality and Intergenerational Links, Review of Economic Studies 71, 743-768.

Del Negro, M., Eggertsson, G., Ferrero, A., and N. Kiyotaki, 2017, The Great Escape? A Quantitative Evaluation of the Fed's Liquidity Facilities, American Economic Review 107, 824-57.

Duffie, Duffie, and Larry G. Epstein, 1992, Stochastic Differential Utility, Econometrica, 60, 353-394.

Duffie, Duffie, Pan, Jun, Singleton, Kenneth J., 2000. Transform analysis and asset pricing for affine jump-diffusions. Econometrica 68, 1343–1376.

Epstein, Larry G.; Zin, Stanley E., 1989, Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework, Econometrica 57, 937–969.

Feldmann I, Anja, and Ward Whitt, 1998, Fitting Mixtures of Exponentials to Long-tail Distributions to Analyze Network Performance Models, Performance Evaluation 31, 245-279.

Gabaix, Xavier, 2009, Power Laws in Economics and Finance. Annual Review of Economics 1, 255-294.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, 2016, The Dynamics of Inequality, Econometrica 84, 2071-2111.

Glasserman, Paul, and Kyoung-Kuk Kim, 2010, Moment Explosions and Stationary Distributions in Affine Diffusion Models, Mathematical Finance 20, 1–33

Gouin-Bonenfant, Emilien, and Alexis Akira Toda, 2020, Pareto Extrapolation: An Analytical Framework for Studying Tail Inequality, working paper, UCSD.

Graham, Paul, 2021, How People Get Rich Now?, http://paulgraham.com/richnow.html.

Grey, D.R. 1994, Regular Variation in the Tail Behavior of Solutions of Random Difference Equations, The Annals of Applied Probability 4, 169-83.

Guvenen, Fatih, 2011, Macroeconomics with Heterogeneity: A Practical Guide, Federal Reserve Bank of Richmond's Economic Quarterly 97, 255-326.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song, 2021, What Do Data on Millions of US Workers Reveal about Life-cycle Earnings Dynamics? forthcoming in Econometrica.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, 2009, Quantitative Macroeconomics with Heterogeneous Households, Annual Review of Economics 1, 319–54.

Heyde, C.C., S. G. Kou, 2004, On the Controversy over Tailweight of Distributions, Operations Research Letters 32, 399-408.

Hubmer, Joachim, Per Krusell, and Anthony A Smith, 2021, The Historical Evolution of the Wealth Distribution: A Quantitative-theoretic Investigation. Martin Eichenbaum & Erik Hurst (eds) NBER Macroeconomic Annual 35, 391-455.

Huggett, Mark. 1993. The Risk-free Rate in Heterogeneous-agent Incomplete-insurance Economies, Journal of Economic Dynamics and Control 17, 953-969.

Jena, R., Kim, K., Xing, H. 2012, Long-term and Blow-up Behaviors of Exponential Moments in Multi-dimensional Affine Diffusions, Stochastic Processes and their Applications 122, 2961-2993.

Jin, Peng, Jonas Kremer, and Barbara Rüdiger, 2020, Existence of Limiting Distribution for Affine Processes, Journal of Mathematical Analysis and Applications 486, 1-31.

Jones, Charles, and Jihee Kim, 2018, A Schumpeterian Model of Top Income Inequality, Journal of Political Economy 126, 1785-1826.

Kaplan, Greg, Benjamin Moll, and Giovanni L. Violante, 2018, Monetary Policy According to HANK. American Economic Review 108, 697-743.

Kaplan, Greg, and Giovanni L. Violante, 2014, A Model of the Consumption Response to Fiscal Stimulus Payments, Econometrica 82, 1199–1239.

Kaplan, Greg, and Giovanni L. Violante, 2021, The Marginal Propensity to Consume in Heterogeneous-Agent Models, in pareparation for Annual Review of Economics.

Keller-Ressel, Martin, and Eberhard Mayerhofer, 2015, Exponential Moments of Affine Processes, Annals Applied Probability 25, 714-752.

Kesten, Harry, 1973, Random Difference Equations and Renewal Theory for Products of Random Matrices." Acta Mathematica 131, 207–48.

Krishnamurthy, A., and Vissing-Jorgensen, A, 2012, The Aggregate Demand for Treasury Debt, Journal of Political Economy 120, 233-267.

Krueger, Dirk, Kurt Mitman, and Fabrizio Perri, 2016, Macroeconomics and Household Heterogeneity, In *Handbook of Macroeconomics*, John B. Taylor and Harald Uhlig (eds), Vol 2A, Elsevier, Chapter 11, 842-921.

Krusell, Per and Anthony Jr. Smith, 1998, Income and Wealth Heterogeneity in the Macroeconomy. Journal of Political Economy 106, 867-896.

Merton, R.C., 1971. Optimal consumption and portfolio rules in a continuous-time model. Journal of Economic Theory 3, 373–413.

Moll, Benjamin, Lukasz Rachel, and Pascual Restrepo, 2019, Uneven Growth: Automations Impact on Income and Wealth Inequality, working paper, Boston University.

Nirei, Makoto and Shuhei Aoki, 2016, Pareto Distribution of Income in Neoclassical Growth Models, Review of Economic Dynamics 20, 25-42.

Quadrini, Vincenzo, 2000, Entrepreneurship, Saving and Social Mobility. Review of Economic Dynamics 3, 1-40.

Quadrini, Vincenzo, and José-Víctor Ríos-Rull, 2015, Inequality in Macroeconomics, In Handbook of Income Distribution, Vol 2, Anthony B. Atkinson and François Bourguignon (eds), Elsevier, 1229-1302.

Saporta, Benoite De, and Yao, Jian-Feng, 2005, Tail of a Linear Diffusion with Markov Switching, Annals of Applied Probability 15, 992–1018.

Sargent, Thomas, Neng Wang, and Jinqiang Yang, 2021, Stochastic Earnings Growth and Equilibrium Wealth Distributions, Tech. Rep. w28473, National Bureau of Economic Research.

Smith, Matthew, Owen Zidar, and Eric Zwick, 2020, Top Wealth in America: New Estimates and Implications for Taxing the Rich, Working Papers 264, Princeton University, Department of Economics, Center for Economic Policy Studies.

Stachurski, John and Alexis Akira Toda. 2019. An Impossibility Theorem for Wealth in Heterogeneous-agent Models with Limited Heterogeneity. Journal of Economic Theory 182 (C), 1-24.

Toda, Alexis Akira, 2014, Incomplete Market Dynamics and Cross-sectional Distributions. Journal of Economic Theory 154, 310-348.

Uzawa, H. , 1968, Time Preference, the Consumption Function, and Optimal Asset Holdings. In J. N. Wolfe (Ed.), Value, Capital, and Growth: Papers in honour of Sir John Hicks, pp. 485-504. Edinburgh: Edinburgh University Press.

Wang, Neng, 2007, An Equilibrium Model of Wealth Distribution, Journal of Monetary Economics 54, 1882–1904

Weil, Philippe, 1993, Precautionary Savings and the Permanent Income Hypothesis, Review of Economic Studies 60, 367–383.

Proofs

Proof of Proposition 1: Let the value function V_t satisfy

$$dV_t = \mu_t dt + \sigma_t^W dW_t^l + \sigma_t^J dN_t,$$

where N_t is a Poisson process with intensity $\lambda_t = \lambda_k k_t$. By Appendix B, the HJB equation is given by

$$\beta f(V_t) = \max_{c_t, k_t} f(c_t) + f'(V_t)\left[\mu_t + \frac{1}{2}\frac{u''(V_t)}{u'(V_t)}\sigma_t^W\left(\sigma_t^W\right)' + \lambda_t \frac{\mathbb{E}_\nu\left(u\left(V_t + \sigma_t^J\right) - u(V_t)\right)}{u'(V_t)}\right], \tag{45}$$

Conjecture that the value function takes the form:

$$V_t = V(x_t, \ell_t) = \theta\left(x_t + \xi_\ell \ell_t + \xi_0\right), \tag{46}$$

where θ, ξ_ℓ, and ξ_0 are constants to be determined. By Ito's Lemma, it follows from (66) and (67) that

$$\begin{aligned} \mu_t &= \theta r x_t + \theta\left(R^k - r\right)k_t - \frac{\theta\eta}{2}k_t^2 - \theta c_t + \theta w \ell_t + \theta\xi_\ell\rho_\ell\left(L - \ell_t\right), &(47)\\ \sigma_t^W &= \theta\xi_\ell\sigma_\ell\sqrt{\ell_t}, \;\; \sigma_t^J = \theta q_t. &(48)\end{aligned}$$

Plugging the above equations into (45) and taking first-order conditions, we have

$$f'(c_t) = \theta f'(V_t), \tag{49}$$

$$\theta\left(R^k - r\right) - \theta\eta k_t + \frac{\lambda_k \mathbb{E}_\nu\left(u\left(V_t + \sigma_t^J\right) - u(V_t)\right)}{u'(V_t)} = 0. \tag{50}$$

By (2) and (46), we have

$$c_t = \theta^{-\psi}V_t = \theta^{1-\psi}\left(x_t + \xi_\ell \ell_t + \xi_0\right), \tag{51}$$

$$k_t = \frac{R^k - r}{\eta} + \frac{\lambda_k \mathbb{E}_\nu\left[1 - \exp\left(-\gamma\theta q\right)\right]}{\eta\gamma\theta}. \tag{52}$$

We then obtain (6) and (7).

We now verify the conjecture (46) and derive coefficients. Using (2) and (51), we can divide the two sides of (45) by $f'(V_t)$ and derive

$$\frac{\beta V_t}{1-1/\psi}=\frac{\theta^{1-\psi}V_t}{1-1/\psi}+\mu_t-\frac{\gamma}{2}\sigma_t^W\left(\sigma_t^W\right)'-\frac{\lambda_t}{\gamma}\mathbb{E}_\nu\left[\exp\left(-\gamma\theta q\right)-1\right].$$

Plugging (47), (48), (46), and (51) into the preceding equation, we obtain

$$\begin{aligned}&\frac{\left(\beta-\theta^{1-\psi}\right)}{1-1/\psi}\theta\left(x_t+\xi_\ell\ell_t+\xi_0\right)\\ =\;&\theta r x_t+\theta\left(R^k-r\right)k_t-\frac{\theta\eta}{2}k_t^2+\theta w\ell_t+\theta\xi_\ell\rho_\ell\left(L-\ell_t\right)\\ &-\theta\theta^{1-\psi}\left(x_t+\xi_\ell\ell_t+\xi_0\right)\\ &-\frac{\gamma}{2}\left(\theta\xi_\ell\sigma_\ell\right)^2\ell_t-\frac{\lambda_k k_t}{\gamma}\mathbb{E}_\nu\left[\exp\left(-\gamma\theta q\right)-1\right].\end{aligned}$$

Matching coefficients yields

$$\frac{\left(\beta-\theta^{1-\psi}\right)}{1-1/\psi}=r-\theta^{1-\psi},\tag{53}$$

$$\frac{\left(\beta-\theta^{1-\psi}\right)}{1-1/\psi}\xi_\ell=w-\left[\rho_\ell+\theta^{1-\psi}\right]\xi_\ell-\frac{\gamma}{2}\theta\left(\xi_\ell\sigma_\ell\right)^2,\tag{54}$$

$$\begin{aligned}\frac{\left(\beta-\theta^{1-\psi}\right)}{1-1/\psi}\xi_0\;=\;&\left(R^k-r\right)k_t-\frac{\eta}{2}k_t^2-\theta^{1-\psi}\xi_0\\ &+\xi_\ell\rho_\ell L-\frac{\lambda_t}{\gamma\theta}\mathbb{E}_\nu\left[\exp\left(-\gamma\theta q\right)-1\right].\end{aligned}\tag{55}$$

Equation (53) gives (9). Simplifying equation (54) gives a quadratic equation

$$\frac{\gamma\theta}{2}\left(\xi_\ell\sigma_\ell\right)^2+\left(r+\rho_\ell\right)\xi_\ell-w=0.$$

The unique positive root gives (10). Equation (55) gives (11). Q.E.D.

Proof of Proposition 2: Optimal wealth x_t satisfies (26). Solving this equation yields

$$x_t=x_0e^{-\rho_x t}+\mu_x\frac{1-e^{-\rho_x t}}{\rho_x}+\int_0^t e^{-\rho_x(t-s)}dJ_s+\int_0^t e^{-\rho_x(t-s)}\phi w\ell_s ds.$$

Given the square-root process (4) and $2\rho_\ell L\geq\sigma_\ell^2$, we have $\ell_t>0$ for all t. Since J_t only jumps upward, we have $\int_0^t e^{-\rho_x(t-s)}dJ_s>0$. We deduce that the support of the long-run stationary distribution of x_t is $(\mu_x/\rho_x,+\infty)$. Thus the result follows from (6). Q.E.D.

Proof of Proposition 3: By assumption $\mathbb{E}_\nu\ln(1+q)<\infty$, it follows from Jin, Kremer, and Rüdiger (2020) that the joint process $\{x_t,z_t\}$ has a stationary distribution. Let $D_t(m,n)$ denote the drift of the process $\widetilde{x}_t^m \widetilde{z}_t^n$. The other part of $\widetilde{x}_t^m \widetilde{z}_t^n$ is martingale terms. Applying Ito's Lemma, we can derive

$$\begin{aligned} D_t(m,n) &= m\widetilde{z}_t^n \widetilde{x}_t^{m-1}\left[-\rho_x \widetilde{x}_t + \phi \widetilde{z}_t - (\lambda_k K)\mathbb{E}_\nu(q)\right] \\ &\quad -\rho_\ell n \widetilde{z}_t^n \widetilde{x}_t^m + \frac{1}{2}n(n-1)\widetilde{x}_t^m \widetilde{z}_t^{n-2}(\widetilde{z}_t + Z)\sigma_z^2 + (\lambda_k K)\widetilde{z}_t^n \mathbb{E}_\nu\left[(\widetilde{x}_t + q)^m - \widetilde{x}^m\right]. \end{aligned}$$

Simplifying yields

$$\begin{aligned} D_t(m,n) &= -\kappa_{m,n}\left(\widetilde{x}_t^m \widetilde{z}_t^n\right) + \frac{1}{2}\sigma_z^2 n(n-1)\left(\widetilde{x}_t^m \widetilde{z}_t^{n-1}\right) + \frac{1}{2}\sigma_z^2 n(n-1)Z\left(\widetilde{x}_t^m \widetilde{z}_t^{n-2}\right) \\ &\quad +\phi m\left(\widetilde{x}_t^{m-1}\widetilde{z}_t^{n+1}\right) - m\widetilde{z}_t^n \widetilde{x}_t^{m-1}(\lambda_k K)\mathbb{E}_\nu(q) + (\lambda_k K)\widetilde{z}_t^n \mathbb{E}_\nu\left[(\widetilde{x}_t + q)^m - \widetilde{x}_t^m\right], \end{aligned}$$

where

$$\kappa_{m,n} \equiv \rho_x m + \rho_\ell n > 0, \quad \rho_x = \vartheta - r = \psi(\beta - r) > 0.$$

By the Binomial expansion formula

$$(\widetilde{x}_t + q)^m = \sum_{j=0}^{m}\binom{m}{j}\widetilde{x}_t^{m-j}q^j, \quad \binom{m}{j} = \frac{m!}{j!\,(m-j)!},$$

we have

$$(\widetilde{x}_t + q)^m - \widetilde{x}_t^m = \sum_{j=1}^{m}\binom{m}{j}\widetilde{x}_t^{m-j}q^j.$$

Thus we can derive

$$\begin{aligned} D_t(m,n) &= -\kappa_{m,n}\left(\widetilde{x}_t^m \widetilde{z}_t^n\right) + \frac{1}{2}\sigma_z^2 n(n-1)\left(\widetilde{x}_t^m \widetilde{z}_t^{n-1}\right) + \frac{1}{2}\sigma_z^2 n(n-1)Z\left(\widetilde{x}_t^m \widetilde{z}_t^{n-2}\right) \\ &\quad +\phi m\left(\widetilde{x}_t^{m-1}\widetilde{z}_t^{n+1}\right) + \sum_{j=2}^{m}(\lambda_k K)\binom{m}{j}x_t^{m-j}\widetilde{z}_t^n \zeta_j, \end{aligned}$$

where

$$\zeta_j = \mathbb{E}_\nu\left[q^j\right] > 0 \text{ for } q > 0.$$

Since $\widetilde{x}_t^m \widetilde{z}_t^n$ is a jump-diffusion process with the drift $D_t(m,n)$, we can derive

$$\begin{aligned} \mathbb{E}\left[\widetilde{x}_t^m \widetilde{z}_t^n\right] &= e^{-\kappa_{m,n}t}\mathbb{E}\left[\widetilde{x}_0^m \widetilde{z}_0^n\right] + \int_0^t e^{-\kappa_{m,n}(t-s)}Q(m,n)\,ds \\ &= e^{-\kappa_{m,n}t}\mathbb{E}\left[\widetilde{x}_0^m \widetilde{z}_0^n\right] + \frac{1}{\kappa_{m,n}}\left(1 - e^{-\kappa_{m,n}t}\right)Q(m,n), \end{aligned} \tag{56}$$

where

$$\begin{aligned} Q(m,n) &= P_0(n)\,\mathbb{E}\left[\widetilde{x}_t^m \widetilde{z}_t^{n-1}\right] + P_0(n)\,Z\mathbb{E}\left[\widetilde{x}_t^m \widetilde{z}_t^{n-2}\right] + P_1(m)\,\mathbb{E}\left[\widetilde{x}_t^{m-1}\widetilde{z}_t^{n+1}\right] \\ &\quad +P_2(m)\,\mathbb{E}\left[\widetilde{x}_t^{m-2}\widetilde{z}_t^n\right] + \sum_{j=3}^{m}P_j(m)\,\mathbb{E}\left[\widetilde{x}_t^{m-j}\widetilde{z}_t^n\right] \end{aligned}$$

and

$$\begin{aligned}
P_0(n) &= \frac{1}{2}\sigma_z^2 n(n-1), \\
P_1(m) &= \phi m, \\
P_2(m) &= (\lambda_k K)\binom{m}{2}\zeta_2, \\
P_j(m) &= (\lambda_k K)\binom{m}{j}\zeta_j, \;\; 3 \le j \le m.
\end{aligned}$$

Since $\kappa_{m,n} > 0$, taking limits in (56) as $t \to \infty$ yields

$$\mathbb{E}\left[\widetilde{x}_t^m \widetilde{z}_t^n\right] = \frac{1}{\kappa_{m,n}} Q(m,n).$$

Then we obtain (32).

We can compute moments $M_{m,n}$ recursively. First, we need to initialize the recursion. That is, we need to specify the moments when either $0 \le m < 3$ or $0 \le n < 2$. We have the following results:

$$\begin{aligned}
M_{0,0} &= 1,\; M_{1,0} = 0,\; M_{0,1} = 0, \\
M_{0,n} &= \frac{1}{\kappa_{0,n}}\left[P_0(n) M_{0,n-1} + P_0(n) Z M_{0,n-2}\right] \;\text{ for } n \ge 2, \\
M_{1,1} &= \frac{1}{\kappa_{1,1}} P_1(1) M_{0,2}, \\
M_{1,n} &= \frac{1}{\kappa_{1,n}}\left[P_0(n) M_{1,n-1} + P_0(n) Z M_{1,n-2} + P_1(1) M_{0,n+1}\right] \text{ for } n \ge 2, \\
M_{2,0} &= \frac{1}{\kappa_{2,0}}\left[P_1(2)M_{1,1} + P_2(2)M_{0,0}\right], \\
M_{2,1} &= \frac{1}{\kappa_{2,1}}\left[P_0(1) M_{2,0} + P_1(2) M_{1,2}\right], \\
M_{2,n} &= \frac{1}{\kappa_{2,n}}\left[P_0(n) M_{2,n-1} + P_0(n) Z M_{2,n-2} + P_1(2) M_{1,n+1} + P_2(2) M_{0,n}\right] \text{ for } n \ge 2.
\end{aligned}$$

$$M_{m,0} = \frac{1}{\kappa_{m,0}}\left[P_1(m)M_{m-1,1} + P_2(m)M_{m-2,0} + \sum_{j=3}^{m} P_j(m) M_{m-j,0}\right] \text{ for } m \ge 3,$$

$$M_{m,1} = \frac{1}{\kappa_{m,1}}\left[P_1(m)M_{m-1,2} + P_2(m)M_{m-2,1} + \sum_{j=3}^{m} P_j(m) M_{m-j,1}\right] \text{ for } m \ge 3.$$

Given these values, we can start the recursive iteration for all $m \le j^*$ if $\zeta_j = \mathbb{E}_\nu\left[q^j\right]$ exists for $1 \le j \le j^*$. Whenever $\zeta_j = \mathbb{E}_\nu\left[q^j\right]$ does not exist for $j = j^* + 1$, all moments $M_{m,n}$ do not exist for $m > j^*$. Q.E.D.

Proof of Propositions 4 and 5: The proof consists of the following four steps by repeatedly applying Proposition 3.

Step 1. We can easily derive the moments for the labor income process:

$$
\begin{aligned}
M_{0,2} &= \frac{\sigma_z^2 Z}{2\rho_\ell},\ M_{0,3} = \frac{\sigma_z^2}{\rho_\ell} M_{0,2} = \frac{\sigma_z^4 Z}{2\rho_\ell^2}, \\
M_{0,4} &= \frac{3\sigma_z^4 Z}{4\rho_\ell^3}\left(\rho_\ell Z + \sigma_z^2\right) = \frac{3\sigma_z^6 Z}{4\rho_\ell^3} + 3M_{0,2}^2.
\end{aligned}
$$

The labor income skewness and excess kurtosis are given by

$$
Skew\left[z\right] = \frac{M_{0,3}}{\left(M_{0,2}\right)^{3/2}} = \frac{\sigma_z^2}{\rho_\ell}\left(M_{0,2}\right)^{-1/2} = \left(\frac{2\sigma_z^2}{\rho_\ell Z}\right)^{\frac{1}{2}},\quad Kurt\left[z\right] = \frac{3\sigma_z^2}{\rho_\ell Z}.
$$

Step 2. Applying Proposition 3, we can derive

$$
\begin{aligned}
Var_x &= M_{2,0} = \frac{1}{\kappa_{2,0}}\left[P_1(2)M_{1,1} + P_2(2)M_{0,0}\right] \\
&= \frac{1}{2\rho_x}\left[2\phi M_{1,1} + \left(K\sigma_k\right)^2 + \left(\lambda_k K\right)\zeta_2\right].
\end{aligned}
$$

Since

$$
M_{1,1} = \frac{1}{\kappa_{1,1}} P_1\left(1\right) M_{0,2} = \frac{\phi M_{0,2}}{\rho_x + \rho_\ell},
$$

we have

$$
\frac{Var_x}{Var_z} = \frac{M_{2,0}}{M_{0,2}} = \frac{\phi^2}{\rho_x\left(\rho_x + \rho_\ell\right)} + \frac{\left(\lambda_k K\right)\zeta_2}{2\rho_x M_{0,2}} = \frac{\phi^2}{\rho_x\left(\rho_x + \rho_\ell\right)} + \varpi_1, \tag{57}
$$

where

$$
\varpi_1 = \frac{\left(\lambda_k K\right)\zeta_2}{2\rho_x M_{0,2}} > 0.
$$

The correlation between x_t and z_t is given by

$$
\frac{M_{1,1}}{\sqrt{Var_x}\sqrt{Var_z}} = \frac{\phi}{\rho_x + \rho_\ell}\frac{\sqrt{Var_z}}{\sqrt{Var_x}}.
$$

Step 3. To compute the wealth skewness, we apply Proposition 3 to derive

$$
\begin{aligned}
M_{1,2} &= \frac{1}{\kappa_{1,2}}\left[P_0\left(2\right)M_{1,1} + P_1\left(1\right)M_{0,3}\right] = \frac{1}{\rho_x + 2\rho_\ell}\left[\frac{\sigma_z^2 \phi M_{0,2}}{\rho_x + \rho_\ell} + \frac{\phi\sigma_z^2}{\rho_\ell}M_{0,2}\right] = \frac{\phi M_{0,3}}{\rho_x + \rho_\ell}, \\
M_{2,1} &= \frac{1}{\kappa_{2,1}} P_1\left(2\right) M_{1,2} = \frac{2\phi M_{1,2}}{2\rho_x + \rho_\ell} = \frac{2\phi^2 M_{0,3}}{\left(2\kappa_x + \rho_\ell\right)\left(\kappa + \rho_\ell\right)},
\end{aligned}
$$

$$
\begin{aligned}
M_{3,0} &= \frac{1}{\kappa_{3,0}}\left[P_1(3)M_{2,1} + P_3\left(3\right)M_{0,0}\right] = \frac{P_1(3)M_{2,1} + \lambda_k K\zeta_3}{3\rho_x} \\
&= \frac{2\phi^3 M_{03}}{\rho_x\left(2\,\rho_x + \rho_\ell\right)\left(\rho_x + \rho_\ell\right)} + \frac{\lambda_k K\zeta_3}{3\rho_x}.
\end{aligned}
$$

The wealth skewness is given by

$$
\begin{aligned}
Skew\left[x\right] &= \frac{M_{3,0}}{\left(M_{2,0}\right)^{3/2}} = \frac{2\phi^3 M_{0,3}\left(M_{0,2}\right)^{3/2}}{\rho_x\left(2\rho_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)\left(M_{0,2}\right)^{3/2}\left(M_{2,0}\right)^{3/2}} + \frac{\lambda_k K\zeta_3}{3\kappa_x\left(M_{2,0}\right)^{3/2}} \\
&= \frac{2\phi^3\left(M_{0,2}\right)^{3/2}}{\rho_x\left(2\kappa_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)\left(M_{2,0}\right)^{3/2}} S_y + \frac{\lambda_k K\zeta_3}{3\kappa_x\left(M_{2,0}\right)^{3/2}} \\
&= Skew\left[z\right]\frac{2\sqrt{\rho_x\left(\rho_x+\rho_\ell\right)}}{2\rho_x+\rho_\ell}\left[1+\frac{\left(\lambda_k K\zeta_2\right)\left(\rho_x+\rho_\ell\right)}{2M_{0,2}\phi^2}\right]^{-3/2} + \frac{\lambda_k K\zeta_3}{3\rho_x\left(M_{2,0}\right)^{3/2}}.
\end{aligned}
$$

Step 4. We finally compute the wealth kurtosis. We use Proposition 3 to derive

$$
\begin{aligned}
M_{2,2} &= \frac{1}{\kappa_{2,2}}\left[P_0\left(2\right)M_{2,1}+P_0\left(2\right)ZM_{2,0}+P_1\left(2\right)M_{1,3}+P_2\left(2\right)M_{0,2}\right] \\
&= \frac{1}{2\left(\rho_x+\rho_\ell\right)}\left[\sigma_z^2 M_{2,1}+\sigma_z^2 ZM_{2,0}+2\phi M_{1,3}+P_2\left(2\right)M_{0,2}\right],
\end{aligned}
$$

$$
\begin{aligned}
M_{1,3} &= \frac{1}{\rho_x+3\rho_\ell}\left[P_0\left(3\right)M_{1,2}+P_0\left(3\right)ZM_{1,1}+P_1\left(1\right)M_{0,4}\right] \\
&= \frac{1}{\rho_x+3\rho_\ell}\left[3\sigma_z^2 M_{1,2}+3\sigma_z^2 ZM_{1,1}+\phi M_{0,4}\right] \\
&= \frac{1}{\rho_x+3\rho_\ell}\left[\frac{3\sigma_z^2\phi M_{0,3}}{\rho_x+\rho_\ell}+\frac{3\sigma_z^2 Z\phi M_{0,2}}{\rho_x+\rho_\ell}+\phi M_{0,4}\right] \\
&= \frac{\phi}{\rho_x+\rho_\ell}M_{0,4}.
\end{aligned}
$$

Plugging the above expression for $M_{1,3}$ into above equation for $M_{2,2}$ and using Step 1 and (57), we can derive

$$
\begin{aligned}
M_{2,2} &= \frac{1}{2\left(\rho_x+\rho_\ell\right)}\left[\frac{2\phi^2\sigma_z^2 M_{0,3}}{\left(2\rho_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)}+\sigma_z^2 ZM_{2,0}+\frac{2\phi^2}{\rho_x+\rho_\ell}M_{0,4}\right]+\frac{P_2\left(2\right)M_{0,2}}{2\left(\rho_x+\rho_\ell\right)} \\
&= \frac{1}{2\left(\rho_x+\rho_\ell\right)}\left[\frac{\phi^2\sigma_z^6 Z}{\rho_\ell^2\left(2\rho_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)}+\sigma_z^2 ZM_{2,0}+\frac{2\phi^2}{\rho_x+\rho_\ell}\left(\frac{3\sigma_z^6 Z}{4\rho_\ell^3}+3M_{0,2}^2\right)\right] \\
&\quad +\frac{P_2\left(2\right)M_{0,2}}{2\left(\rho_x+\rho_\ell\right)}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{1}{2\left(\rho_x+\rho_\ell\right)}\left[\frac{\phi^2\sigma_z^6 Z}{\rho_\ell^2\left(2\rho_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)}+2\rho_\ell M_{0,2}M_{2,0}+\frac{2\phi^2}{\rho_x+\rho_\ell}\left(\frac{3\sigma_z^6 Z}{4\rho_\ell^3}+3M_{0,2}^2\right)\right] \\
&\quad +\frac{\sigma_z^2 Z\varpi_1 M_{0,2}}{2\left(\rho_x+\rho_\ell\right)}
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \frac{1}{2\left(\rho_x+\rho_\ell\right)}\left[\frac{\phi^2\sigma_z^6 Z\left(5\rho_\ell+6\rho_x\right)}{2\rho_\ell^3\left(2\rho_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)}+\frac{2M_{2,0}^2\rho_x\left(\rho_x+\rho_\ell\right)\left(\rho_\ell\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]+3\phi^2\rho_x\right)}{\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]^2}\right] \\
& +\frac{\sigma_z^2 Z\varpi_1 M_{0,2}}{2\left(\rho_x+\rho_\ell\right)},
\end{aligned}
$$

where we have used (57) to derived the last equality.

By Proposition 3, we can compute

$$
\begin{aligned}
M_{4,0} &= \frac{1}{\kappa_{4,0}}\left[P_1(4)M_{3,1}+P_2(4)M_{2,0}+P_4\left(4\right)\right] \\
&= \frac{1}{4\rho_x}\left[4\phi M_{3,1}+P_2(4)M_{2,0}+P_4\left(4\right)\right] \\
&= \frac{\phi\left[3\phi M_{2,2}+P_2\left(3\right)M_{1,1}\right]}{\rho_x\left(3\rho_x+\rho_\ell\right)}+\frac{1}{4\rho_x}\left[P_2(4)M_{2,0}+P_4\left(4\right)\right] \\
&= \frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}M_{2,2}+\varpi_2,
\end{aligned}
$$

where we define

$$
\varpi_2 \equiv \frac{3\phi^2\lambda_k K M_{0,2}\zeta_2}{\rho_x\left(3\rho_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)}+\frac{\lambda_k K}{4\rho_x}\left(6M_{2,0}\zeta_2+\zeta_4\right)>0.
$$

Using the above equation for $M_{2,2}$, we have

$$
\begin{aligned}
M_{4,0} &= \frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}M_{2,2}+\varpi_2 \\
&= \frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}\frac{\phi^2\sigma_z^6 Z\left(5\rho_\ell+6\rho_x\right)}{4\rho_\ell^3\left(2\rho_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)^2} \\
&\quad +3M_{2,0}^2\frac{\phi^2\left(\rho_\ell\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]+3\phi^2\rho_x\right)}{\left(3\rho_x+\rho_\ell\right)\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]^2} \\
&\quad +\frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}\frac{\sigma_z^2 Z\varpi_1 M_{0,2}}{2\left(\rho_x+\rho_\ell\right)}+\varpi_2.
\end{aligned}
$$

We can now compute the wealth excess kurtosis

$$
\begin{aligned}
Kurt\left[x\right] &= \frac{M_{4,0}}{M_{2,0}^2}-3=\frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}\frac{\phi^2\sigma_z^6 Z\left(5\rho_\ell+6\rho_x\right)}{4\rho_\ell^2\left(2\kappa_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)^2}\frac{1}{M_{2,0}^2} \\
&\quad +3\frac{\phi^2\left(\rho_\ell\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]+3\phi^2\rho_x\right)}{\left(3\rho_x+\rho_\ell\right)\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]^2}-3 \\
&\quad +\left[\frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}\frac{\sigma_z^2 Z\varpi_1 M_{0,2}}{2\left(\rho_x+\rho_\ell\right)}+\varpi_2\right]\frac{1}{M_{2,0}^2}.
\end{aligned}
$$

Using (57) , we have

$$
\begin{aligned}
Kurt\left[x\right] &= \left(M_{0,2}\frac{\phi^2}{\rho_x\left(\rho_x+\rho_\ell\right)}+\varpi_1 M_{0,2}\right)^{-2}\frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}\frac{\phi^2\sigma_z^6 Z\left(5\rho_\ell+6\rho_x\right)}{4\rho_\ell^3\left(2\rho_x+\rho_\ell\right)\left(\rho_x+\rho_\ell\right)^2}\\
&+3\frac{\phi^2\left(\rho_\ell\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]+3\phi^2\rho_x\right)}{\left(3\rho_x+\rho_\ell\right)\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]^2}-3\\
&+\left[\frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}\frac{\sigma_z^2 Z M_{0,2}}{2\left(\rho_x+\rho_\ell\right)}\varpi_1+\varpi_2\right]\frac{1}{M_{2,0}^2}.
\end{aligned}
$$

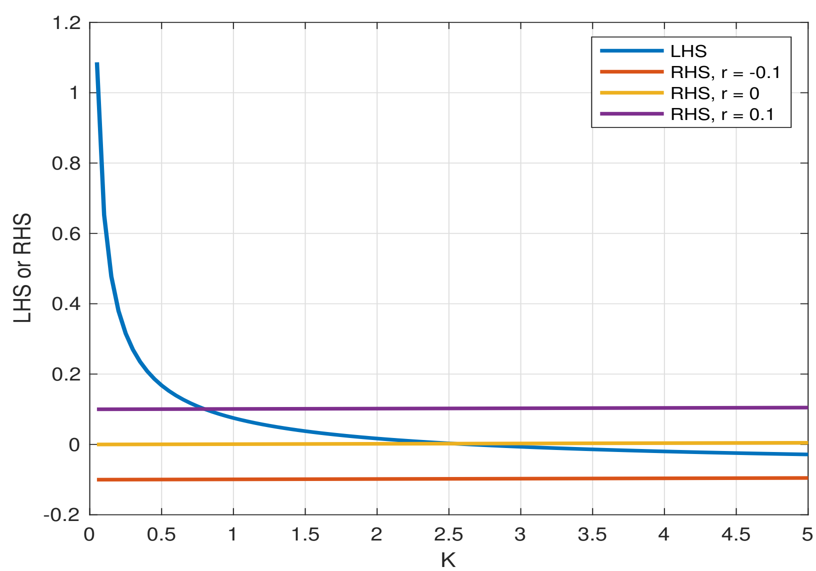
By Step 1, we obtain the wealth excess kurtosis:

$$
\begin{aligned}
Kurt\left[x\right] &= Kurt\left[z\right]\frac{\rho_x\left(5\rho_\ell+6\rho_x\right)}{\left(3\rho_x+\rho_\ell\right)\left(2\rho_x+\rho_\ell\right)}\left(1+\frac{\varpi_1\rho_x\left(\rho_x+\rho_\ell\right)}{\phi^2}\right)^{-2}\\
&+3\frac{\phi^2\left(\rho_\ell\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]+3\phi^2\rho_x\right)}{\left(3\rho_x+\rho_\ell\right)\left[\rho_x\left(\rho_x+\rho_\ell\right)\varpi_1+\phi^2\right]^2}-3\\
&+\left[\frac{3\phi^2}{\rho_x\left(3\rho_x+\rho_\ell\right)}\frac{\sigma_z^2 Z M_{0,2}}{2\left(\rho_x+\rho_\ell\right)}\varpi_1+\varpi_2\right]\frac{1}{M_{2,0}^2}.
\end{aligned}
$$

When $\varpi_1=\varpi_2=0$, the results are reduced to those in Wang (2007).

Proof of Lemma 1: The expression on the left-hand side of (37) decreases with K, goes to $-\chi-\delta$ as $K\to+\infty$, and goes to $+\infty$ as $K\to 0$. The expression on the right-hand side increases with K, goes to $+\infty$ as $K\to+\infty$, and goes to 0 as $K\to 0$. By the intermediate value theorem, there is a unique solution to equation (37). Figure 4 shows one numerical example solution. Since θ increases with r and $\mathbb{E}_\nu\left[1-\exp\left(-\gamma\theta q\right)\right]/\theta$ decreases with θ, we have $\mathbb{E}_\nu\left[1-\exp\left(-\gamma\theta q\right)\right]/\theta$ decreases with r. Thus the line, denoted by RHS in Figure 4, shifts up as r increases. Thus $K\left(r\right)$ decreases with r.

Figure 4: LHS and RHS of (37)



We have $\theta \to \beta^{1/(1-\psi)}$ as $r \to \beta$ and $\theta \to (\psi\beta)^{1/(1-\psi)}$ as $r \to 0$. In both cases, there exists a finite positive solution for K denoted by $K(\beta)$ and $K(0)$. Q.E.D.

Proof of Lemma 2: Plugging (20) into (38) yields

$$\begin{aligned} S &= AK^{\alpha}L^{1-\alpha} + \lambda_k \mathbb{E}_{\nu}[q]K - \vartheta(K + a_h H + \Gamma) - \frac{\eta}{2}K^2 - \chi K \\ &= R^k K + wL + \delta K + \lambda_k \mathbb{E}_{\nu}[q]K - \vartheta(K + a_h H + \Gamma) - \frac{\eta}{2}K^2 - \chi K \\ &= (r + \delta - \vartheta)K + R^k K - rK + wL + \lambda_k \mathbb{E}_{\nu}[q]K - \vartheta(a_h H + \Gamma) - \frac{\eta}{2}K^2 - \chi K. \end{aligned}$$

Using equations (24) and (25), we can derive

$$\begin{aligned} S &= (r + \delta - \vartheta)K + wL(1 - \vartheta a_h/r) \\ &\quad + R^k K - rK + \lambda_k \mathbb{E}_{\nu}[q]K - \vartheta\frac{\eta K^2}{2r} - \frac{\eta}{2}K^2 - \chi K \\ &= (r + \delta - \vartheta)K + wL(1 - \vartheta a_h/r) \\ &\quad + \lambda_k K\left(\mathbb{E}_{\nu}[q] - \frac{\mathbb{E}_{\nu}[1 - \exp(-\gamma\theta q)]}{\gamma\theta}\right) + \frac{1}{2}\eta K^2\left(1 - \frac{\vartheta}{r}\right), \end{aligned}$$

where the last equality follows from (7) and $k = K$.

Now we study the limits. As $r \to 0$, $\vartheta \to \psi\beta$, K tends to a finite limit and hence w tends to a finite limit. As a result, a_h tends to a finite limit in $(0,1)$. We deduce that S tends to $-\infty$.

As $r \to \beta$, we have K tends to a finite limit $K(\beta)$, $\vartheta \to \beta$, and $(1 - \vartheta a_h/r) \to 1 - a_h \in (0,1)$. Since

$$\mathbb{E}_{\nu}[q] - \frac{\mathbb{E}_{\nu}[1 - \exp(-\gamma\theta q)]}{\gamma\theta} > 0,$$

we deduce that S tends to a finite limit, which is larger than $\delta K(\beta)$.

Proof of Proposition 6: By Lemmas 1 and 2 and the intermediate value theorem, there exists a solution $r \in (0,\beta)$ to equation (40). Q.E.D.

Proof of Proposition 7: We have the following affine jump-diffusion process

$$dx_t = -\rho_x x_t dt + \mu_x dt + \phi z_t dt + dJ_t, \tag{58}$$

$$dz_t = \rho_\ell(Z - z_t)dt + \sigma_z\sqrt{z_t}dW_t^\ell, \tag{59}$$

where μ_x is given by (27).

We compute the exponential moment

$$\mathbb{E}\left[\exp(u^x x_t + u^z z_t) \,|\, (x_0, z_0) = (x,z)\right] = \exp(A(t) + B^x(t)x + B^z(t)z), \tag{60}$$

with the boundary conditions $B^x(0) = u^x$, $B^z(0) = u^z$, and $A(0) = 0$, where $u^x > 0$ and $u^z > 0$.

Following Duffie, Pan, and Singleton (2000), we obtain a system of ODEs:

$$\dot{A}(t) = B^x(t)\mu_x + \rho_\ell Z B^z(t) + \lambda_k k \mathbb{E}_{\nu}[\exp(B^x(t)q) - 1], \tag{61}$$

$$\dot{B}^x(t) = -\rho_x B^x(t), \tag{62}$$

$$\dot{B}^z(t) = -\rho_\ell B^z(t) + B^x(t)\phi + \frac{1}{2}(B^z(t)\sigma_z)^2. \tag{63}$$

Then we have $B^x(t) = u^x \exp(-\rho_x t) < u^x$. Given the HED specification for the jump size distribution ν, we have for $u^x < \min_j \{1/\mu_j\}$,

$$\mathbb{E}_\nu \exp(B^x(t) q) < \infty.$$

There are two equilibrium points of the ODE system (62) and (63) for $(B^z(t), B^x(t)) : (0,0)$ and $(2\rho_\ell/\sigma_z^2, 0)$ (see Figure 5). The origin is stable, but the other equilibrium point is unstable. The stability of this system can be analyzed by computing eigenvalues of the linearized system as in Glasserman and Kim (2010). They show that there is a stable set that contains a neighborhood of the origin. The intersection of this region and $\mathbb{R}_+ \times (-\infty, \min_j \{1/\mu_j\})$ gives the stable set for the system (61)-(63). By Keller–Ressell and Mayerhofer (2015), the set of (u^x, u^z) such that

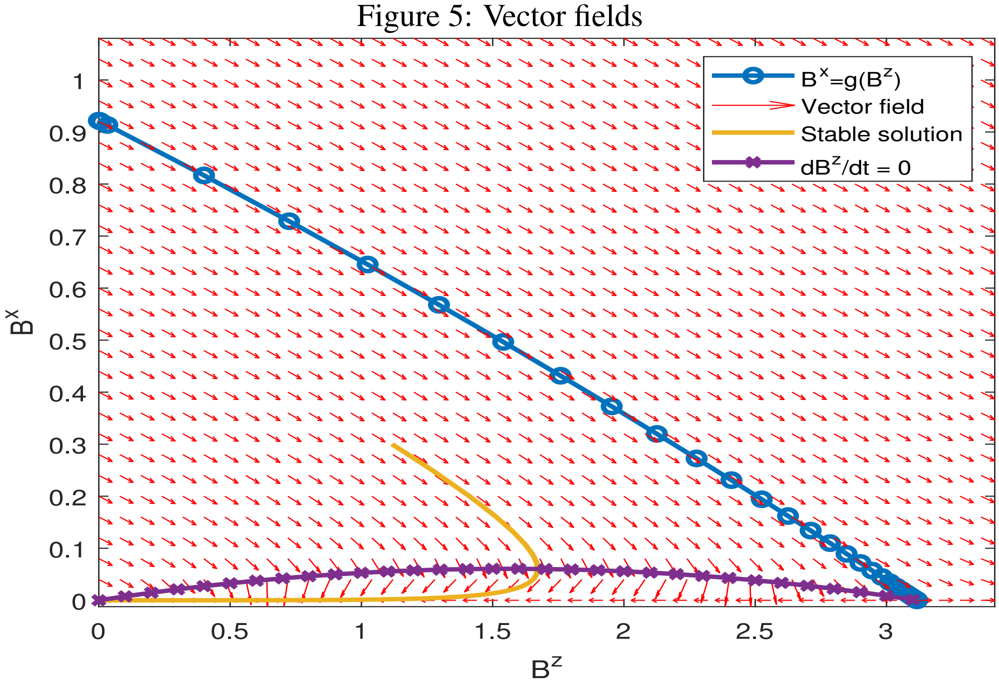
$$\lim_{t\to\infty} \mathbb{E}\left[\exp(u^x x_t + u^z z_t) \,|\, (x_0, z_0) = (x, z)\right] < \infty$$

is the same as the stable set of the system of ODEs (61), (62), and (63). Since this set contains a neighborhood of the origin,

$$\lim_{t\to\infty} \mathbb{E}\left[\exp(\alpha(u^x x_t + u^z z_t)) \,|\, (x_0, z_0) = (x, z)\right]$$

is finite in the set for all $\alpha > 0$ sufficiently small, but it is infinite for $\alpha > 0$ sufficiently large. Thus we conclude that both the stationary distributions of x_t and z_t have an exponential tail to the right by Definition 1. Q.E.D.

Figure 5: Vector fields



Proof of Proposition 8: We continue the analysis in the proof of Proposition 7 by characterizing the stable set more explicitly. By assumption $\phi > 0$ in a stationary equilibrium. We can easily check that the equilibrium point $(0,0)$ is stable and the equilibrium point $(2\rho_\ell/\sigma_z^2,0)$ is a saddle. Moreover, there exists a unique saddle path that converges to the point $(2\rho_\ell/\sigma_z^2,0)$ by inspecting the phase diagram (see Figure 5). The points (B^z,B^x) satisfying $0=-\rho_\ell B^z+B^x\phi+\frac{1}{2}\left(B^z\sigma_z\right)^2$ form the nullcline for $dB^z/dt=0$. The nullcline for $dB^x/dt=0$ is the horizontal line $B^x=0$.

Let $B^x=g\left(B^z\right)$ denote the saddle path in Figure 5. Then for any $B^z\left(0\right)=u^z$, there exists a unique initial value $B^x\left(0\right)=u^x=g\left(u^z\right)$ such that the ODE system (62) and (63) has a unique saddle-path solution for $B^z\left(t\right)$ and $B^x\left(t\right)=g\left(B^z\left(t\right)\right)$ that converges to the equilibrium point $(2\rho_\ell/\sigma_z^2,0)$.

The stable set S_0 for the ODE system for $\left(B^z\left(t\right),B^x\left(t\right)\right)\in\mathbb{R}_+^2$ is given by the region in the figure, whose nonlinear boundary is the saddle path. Let S denote the intersection of S_0 and $\mathbb{R}_+\times\left(-\infty,\min_j\left\{1/\mu_j\right\}\right)$. Note that for $A\left(t\right)$ in (61) to converge, we must have $B^x\left(0\right)=u^x<\min_j\left\{1/\mu_j\right\}$. Then S is the stable set for the ODE system (61), (62), and (63). For any $\left(B^z\left(0\right),B^x\left(0\right)\right)$ in S, we have

$$\lim_{t\to\infty}B^x\left(t\right)=\lim_{t\to\infty}B^z\left(t\right)=0.$$

Moreover,

$$\lim_{t\to\infty}A\left(t\right)=\int_0^\infty\left\{B^x\left(t\right)\mu_x+\rho_\ell ZB^z\left(t\right)+\lambda_k k\mathbb{E}_\nu\left[\exp\left(B^x\left(t\right)q\right)-1\right]\right\}dt$$

is finite by Keller–Ressell and Mayerhofer (2015). Then it follows from their work that

$$\lim_{t\to\infty}\mathbb{E}\left[\exp\left(\alpha\left(u^xx_t+u^zz_t\right)\right)|\left(x_0,z_0\right)=\left(x,z\right)\right]<\infty,$$

for any $\alpha\left(u^z,u^x\right)\in S$ and $\alpha>0$. In particular, we have

$$\lim_{t\to\infty}\mathbb{E}\left[\exp\left(\alpha x_t\right)|\left(x_0,z_0\right)=\left(x,z\right)\right]$$

for any $\alpha\left(0,1\right)\in S$ and $\alpha>0$. Then we can compute the maximum α such that $\alpha\left(0,1\right)\in S$, which gives the exponential decay rate for the stationary distribution of wealth x_t. Since the vertical intercept of the saddle path is equal to $g\left(0\right)$, we have

$$\overline{\alpha}_x=\min\left\{g\left(0\right),\min_j\left\{1/\mu_j\right\}\right\}.$$

Now consider the labor income process z_t. We have

$$\lim_{t\to\infty}\mathbb{E}\left[\exp\left(\alpha z_t\right)|\left(x_0,z_0\right)=\left(x,z\right)\right]<\infty,$$

for $\alpha\left(1,0\right)\in S$ and $\alpha>0$. We can compute the maximum α such that $\alpha\left(1,0\right)\in S$. It satisfies

$$-\rho_\ell\alpha+\frac{1}{2}\left(\alpha\sigma_z\right)^2=0.$$

We then obtain the exponential decay rate for the stationary distribution of labor income z_t, $\overline{\alpha}_z=\frac{2\rho_\ell}{\sigma_z^2}$. Thus, we have $\overline{\alpha}_x<\overline{\alpha}_z$, if

$$\min\left\{g\left(0\right),\ \min_j\left\{1/\mu_j\right\}\right\}<\frac{2\rho_\ell}{\sigma_z^2}.\quad\square$$

Appendix for Online Publication: Continuous-Time Recursive Utility with Jump-Diffusion Risk

We proceed heuristically to derive recursive utility in continuous time by taking limits of a discrete time model (Epstein and Zin (1989) and Duffie and Epstein (1992)). Let dt denote the time increment. The continuation utility U_t at time t satisfies the following recursive equation:

$$U_t = f^{-1}\left[f\left(c_t\right)dt + \exp\left(-\beta dt\right) f\left(\mathcal{R}_t\left(U_{t+dt}\right)\right)\right], \tag{64}$$

where f denotes a time aggregator and $\mathcal{R}_t$ denotes the conditional certainty equivalent.

Suppose that the continuation utility U_t in continuous time satisfies the following backward stochastic differential equation

$$dU_t = \mu_t dt + \sigma_t^W dW_t + \sigma_t^J dN_t, \tag{65}$$

where W_t is a multi-dimensional standard Brownian motion and N_t is a Poisson process with intensity λ_t. The drift μ_t and volatility $\left(\sigma_t^W, \sigma_t^J\right)$ can be derived given a Markovian structure. For example, if U_t depends on the state vector (x_t, ℓ_t) as in our model, denoted by $U_t = U(x_t, \ell_t)$ for a smooth function U, we can apply Ito's Lemma to derive

$$\begin{aligned}
\mu_t &= U_x\left(x_t,\ell_t\right)\left[r x_t + \left(R^k - r\right)k_t + w\ell_t - c_t\right] \qquad (66)\\
&\quad + \frac{1}{2}U_{xx}\left(x_t,\ell_t\right)\left(k_t\sigma_k\right)^2 + U_\ell\left(x_t,\ell_t\right)\rho_\ell\left(L-\ell_t\right) + \frac{1}{2}U_{\ell\ell}\left(x_t,\ell_t\right)\sigma_\ell^2\ell_t,\\
\sigma_t^W &= U_\ell\left(x_t,\ell_t\right)\sigma_\ell\sqrt{\ell_t},\ \ \sigma_t^J = U\left(x_t+q_t,\ell_t\right) - U\left(x_t,\ell_t\right). \qquad (67)
\end{aligned}$$

Given a small time increment dt, we heuristically write $dU_t = U_{t+dt} - U_t$. Then we rewrite (65) as

$$U_{t+dt} = U_t + \mu_t dt + \sigma_t^W dW_t + \sigma_t^J dN_t.$$

By Ito's Lemma, we heuristically write

$$\begin{aligned}
du\left(U_t\right) &= u\left(U_{t+dt}\right) - u\left(U_t\right) = u'\left(U_t\right)\left(\mu_t dt + \sigma_t^W dW_t\right)\\
&\quad + \frac{1}{2}u''\left(U_t\right)\sigma_t^W\left(\sigma_t^W\right)' dt + \left(u\left(U_t + \sigma_t^J\right) - u\left(U_t\right)\right)dN_t.
\end{aligned}$$

Taking conditional expectations yields

$$\begin{aligned}
\mathbb{E}_t u\left(U_{t+dt}\right) &= u\left(U_t\right) + u'\left(U_t\right)\mu_t dt + \frac{1}{2}u''\left(U_t\right)\sigma_t^W\left(\sigma_t^W\right)' dt\\
&\quad + \lambda_t\mathbb{E}_\nu\left(u\left(U_t + \sigma_t^J\right) - u\left(U_t\right)\right)dt,
\end{aligned}$$

where $\lambda_t = \lambda_k k_t$.

Applying a first-order Taylor expansion with respect to dt around zero gives

$$\begin{aligned}
\mathcal{R}_t\left(U_{t+dt}\right) &= u_t^{-1}\mathbb{E}_t u\left(U_{t+dt}\right) = U_t + \mu_t dt + \frac{1}{2}\frac{u''\left(U_t\right)}{u'\left(U_t\right)}\sigma_t^W\left(\sigma_t^W\right)' dt\\
&\quad + \lambda_t\frac{\mathbb{E}_\nu\left(u\left(U_t + \sigma_t^J\right) - u\left(U_t\right)\right)}{u'\left(U_t\right)}dt.
\end{aligned}$$

Following the same procedure again gives

$$f\left(\mathcal{R}_t\left(U_{t+dt}\right)\right)=f\left(U_t\right)+f'\left(U_t\right)\left[\mu_t+\frac{1}{2}\frac{u''\left(U_t\right)}{u'\left(U_t\right)}\sigma_t^W\left(\sigma_t^W\right)'+\lambda_t\frac{\mathbb{E}_\nu\left(u\left(U_t+\sigma_t^J\right)-u\left(U_t\right)\right)}{u'\left(U_t\right)}\right]dt. \quad (68)$$

Up to a first-order approximation, we have

$$\exp\left(-\beta dt\right)=1-\beta dt. \quad (69)$$

Plugging (68) and (69) into (64), we can derive

$$\begin{aligned} f\left(U_t\right) &= f\left(c_t\right)dt+f\left(U_t\right)-\beta\left(c_t\right)f\left(U_t\right)dt \\ &\quad +f'\left(U_t\right)\left[\mu_t+\frac{1}{2}\frac{u''\left(U_t\right)}{u'\left(U_t\right)}\sigma_t^W\left(\sigma_t^W\right)'+\lambda_t\frac{\mathbb{E}_\nu\left(u\left(U_t+\sigma_t^J\right)-u\left(U_t\right)\right)}{u'\left(U_t\right)}\right]dt. \end{aligned}$$

Simplifying yields the continuous-time recursive utility under jump-diffusion uncertainty:

$$\beta f\left(U_t\right)=f\left(c_t\right)+f'\left(U_t\right)\left[\mu_t+\frac{1}{2}\frac{u''\left(U_t\right)}{u'\left(U_t\right)}\sigma_t^W\left(\sigma_t^W\right)'+\lambda_t\frac{\mathbb{E}_\nu\left(u\left(U_t+\sigma_t^J\right)-u\left(U_t\right)\right)}{u'\left(U_t\right)}\right].$$