Capital Income Jumps and Wealth Distribution[†]

Jess Benhabib[‡] Wei Cui[§] Jianjun Miao[¶]

May 6, 2024

Abstract

Compared to the distributions of earnings, 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 jump risks to capital income, which are crucial for generating a thicker tail of the wealth distribution than that of the labor income distribution. Under recursive utility, we derive optimal consumption and wealth in closed form and show that the stationary wealth distribution has an exponential right tail that closely approximates a power-law distribution. Our calibrated model can match the income and wealth distributions in the US data including the extreme right tail of the wealth distribution.

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

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

[†]We are extremely grateful to Stéphane Bonhomme (the editor) and three anonymous referees for insightful comments and suggestions. We thank Steve Kou, Franck Portier, Morten Ravn, Vincent Sterk, Hao Xing, and Shenghao Zhu for helpful discussions. We have also benefited from comments by the participants from many seminars and conferences.

[‡]Dept of Economics, New York University, Email: jess.benhabib@nyu.edu

[§]Dept of Economics, University College London and University of Groningen, 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 (e.g., 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.

In this paper, we provide a tractable general-equilibrium model that accounts for the US distributions of earnings and wealth since 2000, focusing on how sudden new fortunes generated from investments affect the macroeconomy and the wealth distribution. There is ample evidence of sudden new fortunes. For example, by examining 100 of the richest Americans 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." Halvorsen, Hubmer, Ozkan, and Salgado (2023) examine the Norwegian administrative data and find that at least a quarter of the wealthiest (top 0.1%) people start with debt but experience rapid wealth growth early in life as there were some sudden large new fortunes generated from private equity investments.

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 distribution) cannot exceed that of the input (labor income distribution). The reason is that precautionary saving usually compresses the input distribution. By contrast, the capital income jump risks influence precautionary saving in a different way than labor income risks, and therefore our model can generate a thicker tailed wealth distribution than the labor income distribution.

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 maintenance costs (Kaplan and Violante (2014) and Kaplan, Moll, and Violante (2018)). Thus the rate of capital return differs from the interest rate.

Second, we introduce idiosyncratic investment risks in the form of Poisson jumps of capital income from entrepreneurial profits, but not to the rate of return on capital already in place. At each point in time, each household has a chance of investing in a risky project or conducting innovations/R&D. Such activities arrive as rare events and may generate large stochastic profits. These profits are critical to account for the top wealth shares. This feature is consistent with the wealth accumulation of some richest Americans in recent years as mentioned before. Another feature is that the wealth distribution converges quickly since there are always some (albeit very few) people who experience large capital income jumps.[1]

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, Ingersoll, 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 closed-form solution to the individual consumption/saving problem under uninsurable Brownian labor income risk and Poisson capital income jump risk. Unlike the usual exponential-affine models (e.g, Caballero (1990), Calvet (2001), Angeletos and Calvet (2006), and Wang (2003, 2007)), our model setup delivers positive labor income and positive individual consumption under some mild assumptions. The estimated labor income process (with only 3 parameters) matches closely the distribution of income growth obtained from the administrative panel data of Guvenen et al (2021). 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, realistic marginal propensity to consume (MPC) as in the data.[2] 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.[3]

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 income 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. Additionally, our analytical characterization does not rely on permanent earnings heterogeneity which is crucial for deriving analytical results in the literature (e.g., Cao and Luo (2017)), and as mentioned before our earnings process can

[1] In our simulations, the wealth distribution converges to a stationary distribution in less than 15 years in the model.

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

[3] Because we have a closed-form solution for the wealth process, we do not need to use the PDE approach of Achdou et al (2022) to solve for the wealth distribution. As Gouin-Bonenfant and Toda (2023) argue, the usual numerical methods may suffer from large truncation errors at the upper tail of the wealth distribution.

generate a cross-sectional earning-growth distribution similar to the data.

Importantly, having additional capital income jump risks on top of labor earnings risks *does not* guarantee an increase in wealth skewness or in wealth kurtosis. In fact, the precautionary saving motives with jump risks in the BHA environment can even reduce the skewness and the kurtosis of wealth. Our closed-form solution highlights that the illiquidity of capital and higher-order moments of jump risks are critical to increase the skewness and the kurtosis of the wealth distribution. The reason is that they can control the degree of precautionary saving as a result of market incompleteness.

Third, we further provide a novel characterization of the wealth distribution and especially its tail behavior. We show that under our model specification and assumptions, both the stationary wealth and labor income distributions can have a thin exponential tail with all its moments remaining finite if 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 so that the wealth distribution has a heavier tail than the labor income distribution.

To examine the quantitative implications, 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. We adopt the HED specification for the jump size of entrepreneurial capital income because it allows us to get analytic solutions, to compute the stationary equilibrium tractably, and because the HED delivers explicit formulas for the moment generating function and the Laplace transform. Quantitatively, by specifying two exponential components for the HED, we find that our calibrated model can match the wealth distribution in the data closely. In particular, we can match the wealth shares held by the top 0.1% and 1%, despite the fact that the HED and its induced wealth distribution have thin tails because all of their moments remain finite.

Unlike ours, many empirical papers that study the distribution of wealth find that its right tail closely conforms to a power law or a Pareto distribution with a slowly decaying density (see for example Vermeulen (2016) and De Vries and Toda (2021) for extensive studies covering many countries).[4] The recent theoretical literature has also shifted attention to generating a Pareto wealth distribution with thick right tails. This literature applies the Kesten process (Kesten (1973)) with random returns to wealth to give microfoundations to such processes.[5] We refer readers to Gabaix (2009) and Benhabib and Bisin

[4] Vermeulen (2016) computes the tail index for non-response rates for the very rich combining the Forbes 400 list with the Survey of Consumer Finances. He obtains estimates of the Pareto tail index in the range of 1.48-1.55 for the US. Also using the Forbes 400 for the period 1988-2003, Klass et al (2006) estimate a Pareto tail index for the US of 1.49. Clementi and Gallegati (2005) find thick tails of wealth for Italy from 1977 to 2002. Dagsvik et al (2013) use the power law distribution to describe the distribution of wealth in Norway in 1998, while Vermeulen (2016) finds the same for several European countries. Cowell (2011) also finds Pareto tail indices of 1.63-1.85 for wealth in Sweden, and 1.33-1.54 in Canada. See also Jones (2015).

[5] See for example 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 (2014, 2019), Sargent, Wang, and Yang (2021), and Moll, Rachel, and Restrepo (2022). See Hubmer, Krusell, and Smith (2021) for a quantitative study.

(2018) for surveys of this literature and additional references. So at a first sight our results may seem quite surprising, while in fact they can be easily reconciled with the empirical and theoretical results just cited. 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) 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. Moreover, the HED can closely approximate any completely monotone distribution including the Pareto distribution (see Feldmann and Whitt (1998)).[6] Given finite wealth distribution data, our results are compatible with Pareto distributed wealth, and can be empirically well-approximated either with an appropriate power law distribution, or with a thin tailed wealth distribution that is generated by an appropriate thin tailed HED for the jump size of entrepreneurial capital income.

Related literature. Our paper contributes to the macroeconomics literature on wealth inequality in the tradition of the BHA model.[7] Examples include the warm-glow bequest and human capital motives of De Nardi (2004) which can make rich agents save at higher rates, 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), or random idiosyncratic returns to wealth. Our paper is also closely related to the work on the importance of entrepreneurship of Quadrini (2000) and Cagetti and De Nardi (2006).

In an important paper, Stachurski and Toda (2019) prove that if (i) agents are infinitely-lived, (ii) saving is risk-free, and (iii) agents have constant discount factors, then the wealth distribution inherits the tail behavior of income shocks (e.g., light-tailedness or the Pareto exponent). Their results show conclusively that it is necessary to go beyond standard BHA models to explain the empirical fact that wealth is heavier-tailed than income (see Benhabib and Bisin (2018) for related results).

Our modeling of capital income jumps is similar to that of a rare event of jumping from low earnings to very high earnings in Castaneda, Diaz-Gimenez, and Rios-Rull (2003). Besides many other differences, Castaneda, Diaz-Gimenez, and Rios-Rull (2003) assume that both liquid bond assets and illiquid capital assets earn the same rate of return, while we explicitly distinguish between these two types of assets and assume that jumps happen to the capital income only, but not the labor income. Moreover, the jump intensity is endogenous in our model, while it is exogenous in their model.

Our modeling of investment risks builds on the early work of Quadrini (2000) on stochastic arrivals of entrepreneurial profit opportunities to a random subset of households each period, as well as the work of Cagetti and De Nardi (2006), Angeletos and Calvet (2006), Angeletos (2007), and Angeletos and Panousi (2011). Our model differs from these papers in three features, in addition to some technical details: (i) We introduce capital income jump risks (instead of Brownian risks) in addition to labor income risks. (ii) We adopt the recursive utility specification of Weil (1993) in a continuous-time setup, which ensures

[6] 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 $n \geq 1$.

[7] See Heathcote, Storesletten, and Violante (2009), Guvenen (2011), Quadrini and Rios-Rull (2015), Krueger, Mitman, and Perri (2016), and Benhabib and Bisin (2018) for recent surveys.

consumption is always positive under mild conditions. To the best of our knowledge, our paper is the first one to embed the Weil's discrete-time model in a continuous-time setup. (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)) and a jump-diffusion labor income process in an endowment economy. 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. His model implies counterfactually that wealth is less skewed and is lighter tailed than income. By contrast, our recursive utility specification and the new ingredient of capital income jumps allow us to generate realistic MPC and 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 efficiency labor units 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.[8]

For simplicity, we omit the household-specific index i for now. 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 recursion:

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

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 $f(U_t)$. We

[8]See Caldara et al (2012) for a comparison of different solution methods for computing the dynamic stochastic general equilibrium models with recursive preferences in discrete time.

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 closed-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 a consumption-dependent rate of time preference (Uzawa (1968)).[9] But his model still generates negative consumption.

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. The government imposes a tax rate τ_k on the capital income and a tax rate τ_ℓ on the labor income. 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 aggregate productivity, output, capital, and labor, respectively. Profit maximization implies

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

where w and R^k denote the *after-tax* 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

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

market labor efficiency units are represented by the process (ℓ_t), which is governed 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 $\sigma_\ell, \rho_\ell > 0$. This is the square-root process modeled in Cox et al. (1985). One can interpret ℓ_t as the product of labor hours and idiosyncratic labor productivity. To ensure ℓ_t is strictly positive for theoretical proofs, we assume that $2\rho_\ell L \geq \sigma_\ell^2$. The process has a time-varying volatility and the long-run mean of ℓ_t is equal to the aggregate labor supply L.[10]

To model the investment risk, suppose that the capital income from entrepreneurial profits is hit by a jump shock dJ_t, where J_t is a jump process. Notice that this income generates another form of output from business risk-taking. For each realized jump, the *random* 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. For notation simplicity, q already takes into account the tax rate τ_k; the before-tax jump size is $q/(1 - \tau_k)$.

We assume 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 new investment earns a large return and the success probability is positively related to invested capital. Such a return represents additional output from entrepreneurial risk-taking activities like initiating new projects, innovations, or R&D.[11] 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). In our model, investment returns are additive to labor earnings, and may be viewed as high entrepreneurial earnings in excess of labor earnings. This is very similar to the awesome states or rare events in which individual labor productivity can become extremely high, as in Castaneda, Diaz-Gimenez, and Rios-Rull (2003)). The key difference is that in our model the large and rare income comes from entrepreneurial investment only, instead of labor income, and capital earns a higher expected rate of return than the interest rate on bonds. Moreover, the jump intensity $\lambda_t = \lambda_k k_t$ is endogenous in our model, while it is exogenously fixed in their model.

Capital assets are illiquid and owning k_t of them incurs maintenance costs given by $\eta k_t^2/2 + \chi k_t$

[10]The discretized version of (4) is

$$\ell_{t+dt} - \ell_t = \rho_\ell (L - \ell_t)\, dt + \sigma \sqrt{\ell_t dt} \epsilon_{t+dt},$$

where dt is the time increment and ϵ_{t+dt} is an independent standard normal random variable. This is a type of the ARCH model in that the variance of ℓ_{t+dt} conditional on ℓ_t is $\sigma^2 \ell_t dt$. Related models are used to estimate earnings dynamics in the literature, e.g., Arellano et al (2022).

[11]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.

per unit of time, where $\eta > 0$ and $\chi > 0$ are parameters. The household can trade riskless bonds at the after-tax interest rate r to insure against idiosyncratic shocks. Let b_t denote the household's holding of private and public bonds. Households can borrow and lend among themselves without any trading frictions so that $b_t < 0$ represents borrowing. To deliver a closed-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)).

The entrepreneurial profits (capital income) π_t follow dynamics

$$d\pi_t = R^k k_t dt - \left(\chi k_t + \frac{\eta}{2} k_t^2\right) dt + dJ_t.$$

By contrast, the continuous-time literature on investment risks (e.g., Angeletos and Panousi (2011)) typically focuses on Brownian shocks and replaces dJ_t with $\sigma_k k_t dB_t^k$, where B_t^k is another Brownian motion. Let $x_t = b_t + k_t$ denote the household's wealth level. Then the household faces the following dynamic budget constraint

$$dx_t = rb_t dt + d\pi_t + w\ell_t dt - c_t dt + \Upsilon dt,$$

where Υ represents per capita government transfers (or lump-sum taxes if $\Upsilon < 0$). Combining the above two equations yields the wealth dynamics:

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

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 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

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

where θ, ξ_ℓ, and ξ_0 are given by

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

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

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

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 (after-tax) 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 maintenance cost, which reflects the 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.[12]

To understand the consumption rule in (6), we need to introduce the concept of human wealth, which is defined as the (after-tax) 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

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

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

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

[12] One can obtain many different constant levels of capital investment by, e.g., introducing different values of EIS ψ to various groups of households in the economy. Then rich households who hold more capital will have even higher chances to generate capital income jumps, making the jumps even more powerful in generating the wealth dispersion. We choose to be conservative by having only one single level of capital, which also simplifies the numerical exercise later.

The consumption rule in (14) is related to the permanent income theory of consumption. It implies that optimal consumption is linear in human and nonhuman wealth.[13] This property is important for aggregation and useful to analyze the wealth distribution and stationary equilibrium. 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)) without binding borrowing constraints. Importantly, recursive utility in our model helps generate a MPC higher than r.

It follows from (10) that $a_h \in (0, 1)$ (see the proof of Proposition 1 in Appendix A). 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-income levels. 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 $(1 - a_h) h_t$, which is stochastic. The remaining term in (14), $\Gamma = \eta k^2 / (2r) + \Upsilon / r$, can be rewritten as

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

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 saving against capital income jump risks.

2.3 Government

Let G be the exogenous government expenditure at each time. The government has an exogenous fixed bond supply B at each time. As mentioned above, the government tax capital income and labor income at flat rate τ_k and τ_ℓ, respectively. The residual Υ is used as lump-sum transfer. When $\Upsilon < 0$, it becomes lump-sum tax. In a stationary equilibrium, the government budget constraint is given by

$$G + \Upsilon + rB = \frac{\tau_k}{1 - \tau_k} (R^k + \lambda_k \mathbb{E}_\nu [q]) K + \frac{\tau_\ell}{1 - \tau_\ell} wL, \tag{18}$$

where K denotes aggregate capital stock.

[13] See Hall (1978) for an early empirical test of the permanent income theory. See Weil (1993) for a further discussion of the empirical implications of the consumption rule in (14). See Carroll and Kimball (1996) for a discussion of theory and evidence against linearity of the consumption function.

2.4 Stationary Equilibrium

We now add household-specific index i and conduct aggregation. Aggregate consumption, labor, capital, wealth, and output are given by

$$\begin{aligned} C_t &\equiv \int c_t^i di,\ L \equiv \int \ell_t^i di,\ \ K_t \equiv \int k_t^i di,\ X_t \equiv \int x_t^i di, \\ Y_t &\equiv \int y_t^i di + \frac{\lambda_k}{1-\tau_k}\mathbb{E}_\nu[q] \int k_t^i di = AK_t^\alpha L^{1-\alpha} + \frac{\lambda_k}{1-\tau_k}\mathbb{E}_\nu[q] K_t. \end{aligned}$$

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. After aggregation, we are ready to define equilibrium in the steady state.

Definition. *Given constant government policy* $(G, B, \tau_k, \tau_\ell)$*, 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* $i \in [0,1]$ *, a transfer policy* Υ*, 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, capital, and labor markets all clear*

$$\int b_t^i di = B,\ X = K + B,\ \int l_t^i di = L,$$

and (iii) the government budget constraint (18) *holds.*

By (14), aggregate consumption is given by

$$C = \vartheta \left(K + B + a_h H + \Gamma\right), \tag{19}$$

where Γ is defined in (17) with $k = K$ and we can write human wealth (from (13)) as

$$H \equiv \int h_t^i di = \frac{wL}{r}. \tag{20}$$

Notice that our model implies a version of aggregate Ricardian equivalence in the following sense. A dollar increase in the government bond supply B can be offset by r dollars decrease in the government transfer. As a result, the changes in B and Γ offset each other so that aggregate consumption in (19) does not change with B. All other aggregate equilibrium variables do not change with B either. However, debt policy B has an impact on the individual decisions and shifts the wealth distribution. Intuitively, every household prepares exactly enough to offset the consequent change in the lump-sum transfer/taxes and bond holdings, and this shift has unequal impacts on households in an environment with uninsurable idiosyncratic risks. We will provide a further discussion on the wealth dynamics later.

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

$$R^k = (1-\tau_k)\left(\alpha A K^{\alpha-1} L^{1-\alpha} - \delta\right), \tag{21}$$
$$w = (1-\tau_\ell)(1-\alpha) A K^{\alpha} L^{-\alpha}, \tag{22}$$

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

$$\frac{R^k}{1-\tau_k}K + \frac{w}{1-\tau_\ell}L = AK^{\alpha}L^{1-\alpha} - \delta K. \tag{23}$$

Aggregating the budget constraints (5) and (18) and using (23) and the market-clearing condition $X = K + B$, we obtain the resource constraint (so that we verify the Walras' law)

$$C + G + \delta K + \frac{\eta}{2}K^2 + \chi K = Y, \tag{24}$$

where aggregate output is given earlier.

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

$$dx_t = -\rho_x x_t dt + \mu_x dt + \phi w \ell_t dt + dJ_t, \tag{25}$$

where

$$\mu_x \equiv \left(R^k - \chi - r\right)k - \frac{\eta}{2}k^2 - \vartheta\xi_0 + \Upsilon, \tag{26}$$

$$\rho_x \equiv \psi(\beta - r), \;\; \phi \equiv 1 - \frac{\vartheta\xi_\ell}{w}. \tag{27}$$

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. Notice further that a dollar increase in bond policy together with an r-dollar reduction in lump-sum transfers does not change the government budget constraint and the aggregate equilibrium variables, but changes μ_x (e.g., because of a change in ξ_0 as a function of lump-sum transfers Υ). Thus the limiting distribution of x_t changes with the debt policy.

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

$$dz_t = \rho_\ell (Z - z_t) dt + \sigma_z \sqrt{z_t} dW_t^l, \tag{28}$$

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 $\mathbb{E}_\nu \ln(1+q) < \infty$ (Jin, Kremer, and Rüdiger (2020)). The assumption on the jump size 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, and 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 $\mathbb{E}_\nu \ln(1+q) < \infty$. 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.[14]

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 income 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 (25) that the demeaned processes satisfy the dynamics

$$d\widetilde{x}_t = [-\rho_x \widetilde{x}_t + \phi \widetilde{z}_t - \lambda_k K \mathbb{E}_\nu (q)] dt + dJ_t, \tag{29}$$

[14] By our calibration in Section 4, this condition is indeed satisfied.

$$d\widetilde{z}_t = -\rho_\ell \widetilde{z}_t dt + \sigma_z \sqrt{\widetilde{z}_t + Z} dW_t^l, \tag{30}$$

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 $\mathbb{E}_\nu \ln(1+q) < \infty$. Suppose that $\zeta_j \equiv \mathbb{E}_\nu [q^j] > 0$ is finite for $1 \le j \le j^*$ 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 $0 \le m \le j^*$ and $n \ge 0$:*

$$M_{m,n} = \frac{1}{\kappa_{m,n}} \left[\begin{array}{c} P_0(n) M_{m,n-1} + P_0(n) Z M_{m,n-2} + P_1(m) M_{m-1,n+1} \\ +P_2(m) M_{m-2,n} + \sum_{j=3}^m P_j(m) M_{m-j,n} \end{array} \right], \tag{31}$$

where $M_{0,0} = 1, M_{0,1} = M_{1,0} = 0$, $\kappa_{m,n} \equiv \rho_x m + \rho_\ell n$,

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

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 $0 \le m < 3$ or $0 \le n < 2$. We provide the details in Appendix A.

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

$$\frac{\text{Var}[x]}{\text{Var}[z]} = \frac{\phi^2}{\rho_x(\rho_x + \rho_\ell)} + \frac{\lambda_k K \zeta_2}{2\rho_x \text{Var}[z]}, \tag{32}$$

where $\text{Var}[z] = \sigma_z^2 Z/(2\rho_\ell)$ is the long-run income variance. The correlation between the wealth and labor income processes is given by

$$\frac{\phi}{\rho_x + \rho_\ell}\sqrt{\frac{\text{Var}[z]}{\text{Var}[x]}}. \tag{33}$$

As equation (25) shows, the variable ϕ can be interpreted as the MPS out of labor income. The first term on the right side of (32) 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 income variability contributes to the wealth variance as well. The second term in (32) reflects this contribution which comes from the random capital income jump. Equation (33) 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 (excess) kurtosis, denoted by $\text{Skew}[x]$ and $\text{Kurt}[x]$ for wealth x_t, and by $\text{Skew}[z]$ and $\text{Kurt}[z]$ for labor income z_t. In Appendix A we derive the following result.

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

$$\text{Skew}[x] = \text{Skew}[z] \frac{2\sqrt{\rho_x(\rho_x+\rho_\ell)}}{2\rho_x+\rho_\ell}\left[1+\frac{(\lambda_k K\zeta_2)(\rho_x+\rho_\ell)}{2M_{0,2}\phi^2}\right]^{-3/2}+\frac{\lambda_k K\zeta_3}{3\rho_x(M_{2,0})^{3/2}}, \tag{34}$$

and

$$\begin{aligned}\text{Kurt}[x] \quad = \quad & \text{Kurt}[z]\frac{\rho_x(5\rho_\ell+6\rho_x)}{(3\rho_x+\rho_\ell)(2\rho_x+\rho_\ell)}\left[1+\frac{\varpi_1\rho_x(\rho_x+\rho_\ell)}{\phi^2}\right]^{-2} \\ & +3\frac{\phi^2(\rho_\ell[\rho_x(\rho_x+\rho_\ell)\varpi_1+\phi^2]+3\phi^2\rho_x)}{(3\rho_x+\rho_\ell)[\rho_x(\rho_x+\rho_\ell)\varpi_1+\phi^2]^2}-3 \\ & +\left[\frac{3\phi^2}{\rho_x(3\rho_x+\rho_\ell)}\frac{\sigma_z^2 Z M_{0,2}\varpi_1}{2(\rho_x+\rho_\ell)}+\varpi_2\right]\frac{1}{M_{2,0}^2},\end{aligned} \tag{35}$$

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 income 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 (34) shows that the capital income jump risk reduces the wealth skewness relative to the earnings skewness. However, the second term raises the wealth skewness when $\zeta_3 > 0$.

Equation (35) shows that the wealth kurtosis consists of three components. The capital income jump risk reduces the first component (the first line of (35)) as $\varpi_1 > 0$. We can also show that the second component (the second line of (35)) is negative. Only the last component (the third line of (35)) 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 income 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 the next subsection we will provide an explicit characterization of the tail behavior of the wealth and labor income distributions.

3.2 Exponential Tail

To provide a sharp characterization of the tail distribution of wealth, we must specify the jump size distribution ν explicitly. 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 distribution is flexible and can approximate any completely monotone distributions (Feldmann and Whitt (1998)).[15] The PDF for the HED can be written as

$$f(q) = \sum_{j=1}^{n} p_j \frac{\exp(-q/\mu_j)}{\mu_j}, \ q > 0, \tag{36}$$

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

Given the HED (36) 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 moment of x_t as $t \to \infty$, $\lim_{t\to\infty} \mathbb{E}[\exp(\alpha x_t)]$ for $\alpha > 0$, following Glasserman and Kim (2010), Jena, Kim, and Xing (2012), and Keller-Ressel and Mayerhofer (2015). Notice that when a limiting stationary distribution exists, the limiting exponential moment of x_t does not depend on the initial value x_0. By definition, if the limiting exponential moment is finite for any $\alpha \in [0, \alpha_0)$ for some $\alpha_0 > 0$, but is infinite for any $\alpha > \alpha_0$, then the stationary wealth distribution has an exponential tail. See Definition 1 in Appendix C for a general definition of tail behavior.

Proposition 6. *Given the HED (36) 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.*

In the macroeconomics literature the wealth and income distributions are often estimated by Pareto or power law distributions with fat right tails. 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. Given the results of Heyde and Kou (2004), our quantitative results in Section 4.2 show that our wealth process can match the wealth distribution data reasonably well, even if the wealth distribution does not have a Pareto tail.

Proposition 6 shows that both wealth and labor income distributions have exponential tails in our

[15] 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.

model. The following proposition characterizes the exponential decay rates of the stationary wealth and labor income distributions. See Appendix C for a discussion of differences from Pareto tails.

Proposition 7. *Suppose that the jump size follows the HED (36) 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*

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

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 $\overline{\alpha}_x = 1/\mu_2 = 0.0024$ and $\overline{\alpha}_z = 0.5546$. 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 the capital income jump size distribution. If the capital income 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 income jumps. Without investment jump risks, the wealth distribution would have a lighter tail than the income distribution (Benhabib and Bisin (2018) and Stachurski and Toda (2019)).

3.3 Aggregate Investment and Saving

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 the stationary equilibrium quantities and prices independent of the full wealth distribution.

We first derive the aggregate investment demand curve. Combining equations (7) and (21) yields

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

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 is a continuous and decreasing function of 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 saving S is given by

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

where we have substituted the aggregate consumption function in (19) into the above equation. Since aggregate capital K is a function of r, aggregate output Y is also a function of r. We use (21) and (22) to derive R^k and w as functions of r and hence aggregate consumption is also a function of r by (17) and (20). As a result, S is a function of r. In Appendix A, we show that

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

Aggregate saving has five components: The first component $(r+\delta-\vartheta)K$ represents saving out of capital assets. The second component is precautionary saving against the Brownian labor income risk. The third component represents saving out of capital returns. The fourth component represents precautionary saving against the capital income jump risk. The last component is proportional to public saving (taxes minus government expenditure excluding lump-sum transfers/taxes).

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 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 saving in (39) approaches $\delta K(\beta)$. But all other components of aggregate saving in (39) are nonnegative and precautionary saving is strictly positive. We thus deduce that aggregate saving exceeds aggregate investment as $r \to \beta$.

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

Proposition 8. *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 income to make the aggregate comparison consistent. 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:

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

The saving function becomes

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

In steady-state equilibrium, we have $S^c(r) = \delta K(r)$, which leads to $r = \beta = \vartheta$. Alternatively, one may follow Aiyagari (1994) and take the saving function as $r = \beta$.

(ii) The BHA economy with only labor income risks, in which capital and bond assets are perfect substitutes by assuming $\eta = \chi = 0$ and each household receives $\lambda_k \mathbb{E}_\nu[q]$ as capital income to make the aggregate consistent. Importantly, total returns on bonds and on capital are identical. In this case, investment demand $\delta K(r)$ is determined by

$$(1-\tau_k)(\alpha A K^{\alpha-1} L^{1-\alpha} - \delta) + \lambda_k \mathbb{E}_\nu[q] = r,$$

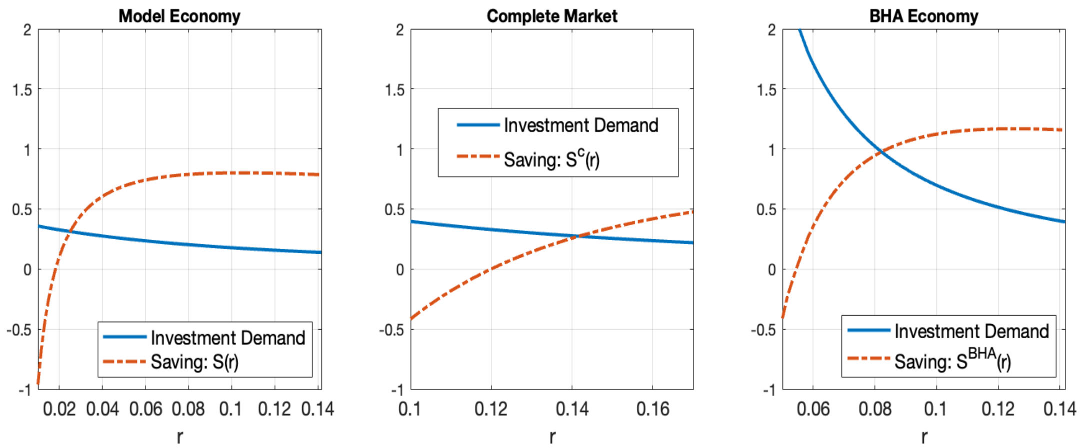
and the saving function is

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

By similar arguments in the proof of Proposition 8, we have $0 < r < \beta$.

Figure 1 illustrates the three economies with the parameter values calibrated in the next section. Notice that the parameter values are not important for the illustrative purpose of our comparison. The investment demand curves are all downward sloping functions of the interest rate r. In the complete markets case, the saving curve intersects 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.

Figure 1: Investment and saving functions in three economies



Notes: We use the parameter values obtained from the calibration in Section 4.1.

We close this section by discussing the equilibrium level of the capital stock. Angeletos (2007) argues that the investment risk raises precautionary saving, 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. Given our calibrated parameter values discussed in the next section, the steady-state capital stock under incomplete markets of our model is slightly higher than that under complete markets. The main reason is that the decrease in the investment demand due to large investment (jump) risks is still dominated by the precautionary-saving effect. However, when our model is compared to the BHA economy (in which we eliminate the capital income risk but keep the labor income risk), the investment demand in our model is much lower, generating a lower level of the steady-state capital stock.

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 $\{\alpha, \delta, \psi, \chi, \eta, \beta, \gamma, A, G, B, \tau_k, \tau_\ell\}$ (see Table 2). We set the capital share $\alpha = 0.33$ as in the macro literature. Set the depreciation rate $\delta = 12\%$ to target 16%

investment to output ratio in the US data. We set the EIS parameter $\psi = 1.5$ in line with the finance literature on long-run risk, and later we conduct a sensitivity analysis with respect to ψ. Set the linear maintenance cost parameter χ, the quadratic maintenance cost parameter η, the subjective discount rate β, and the CARA parameter γ to target the following equilibrium variables: the interest rate $r = 2.5\%$ in line with the real return of government bonds,[16] the equity premium $R^k - r = 3.7\%$,[17] the MPC $\vartheta = 0.20$ by (15), in line with most of OECD aggregate MPC measures (Carroll, Slacalek, and Tokuoka (2014) and Kaplan and Violante (2021)), and the coefficient of relative risk aversion $\gamma C = 5$, where C is the aggregate consumption level in the stationary equilibrium. We normalize the steady-state (after-tax) wage rate to one by adjusting the TFP parameter A.

Government spending G is set so that the government expenditure to output ratio is 19% in line with the data. The debt to output ratio is 81%, which is the average between 2000 and 2016. This period is also used for our distributional statistics discussed later. Since the model does not feature borrowing constraints and the government has the lump-sum tax instrument, calibrating the government debt to a different level has no aggregate consequence because of the Ricardian equivalence discussed before. Both capital tax and labor income tax rates are set to 25%, so the government collects 25% of output as tax revenues.[18]

Earnings Process. Next, we use the simulated method of moments (SMM) to estimate the three parameters L, ρ_ℓ, and σ_ℓ in (4) to target some important statistics in the social security administrative (SSA) data analyzed by Guvenen et al (2021).[19] We focus on the statistics of the changes of log (annual) earnings (i.e., approximately annual earnings growth) after controlling for various factors (e.g., education). Since we do not have corresponding continuous-time earnings data, it is best to explore both the time series and cross-sectional information of the discretized model in (4). The time increment dt corresponds to one week, i.e., $1/52$. Notice that Kaplan et al (2018) specify log earnings as the sum of two jump-drift processes and use the SMM approach to estimate 6 parameters. Guvenen et al (2021) also use this approach to estimate a much more complicated specification that includes a mixture of normal innovations and a nonemployment shock. Our square-root earnings process with only 3 parameters in (4), however, is designed to be simpler for our tractable analysis, but it still captures important statistics of earnings growth as shown in Table 1.

The square-root process has a known stationary distribution, which follows the Gamma type with a

[16] We use the average returns of 1 year Treasury bills and long-term Treasury bonds between 2000 and 2020. The result is robust if we target different maturities, and the target interest rate r is in the range of 2% to 3%.

[17] Our calibrated equity premium is the sum of three components. First, 0.7% of the liquidity premium (because of the maintenance costs of capital), which can be approximated by the average of the spread between AAA corporate bonds and treasuries of similar maturity after 2000 (see Krishnamurthy and Vissing Jorgensen (2012), Del Negro et al (2017), and Cui and Radde (2020)). Second, 2% of the private equity premium for compensating idiosyncratic risks according to Angeletos (2007). Third, 1% of the premium due to other reasons.

[18] We experiment with different tax rates (available upon request) and find that our results are similar after recalibration.

[19] The SSA data cover a long time span from 1978 to 2013, with a substantial sample size (10% random sample of males aged 25-60). Total annual labor earnings are the sum of total annual wage income from W-2 forms and the labor portion (2/3) of self-employment income from Schedule SE.

Table 1: Statistics from the estimated log earnings changes (1 year)

log-earnings changes	Std. Dev.	Skewness	Kurtosis	fraction $< 5\%$	$< 10\%$	$< 20\%$
Data	0.51	- 1.07	14.93	30.6%	48.8%	66.5%
Model	0.51	- 0.002	15.94	28.9%	48.7%	68.7%

Notes: The data statistics are from Figure 1 and Table 1 in Guvenen et al (2021). The model result is based on the SMM estimates $\rho_\ell = 0.0033$, $\sigma_\ell = 0.1098$, and $L = 0.8033$. Each simulation has 10^5 agents who start with levels of earnings drawn from the invariant distribution implied by the earnings process. For each agent, we simulate the earnings for two years and calculate the statistics from the cross-sectional distribution of the log (annual) earnings changes. The calculation is repeated for 100 times and we take the averages. A minimizer routine is implemented to search for the parameters that minimize the distance between the model and the data (excluding kurtosis). Each dt is approximated by one week, which is 1/52, as a year has roughly 52 weeks. Therefore, given parameters, the simulation has $10^5 \times 100 \times 104 = 1.04 \times 10^9$, roughly 1 billion, person-week observation. Thus, the minimizer routine for the SMM requires 1 billion person-week observations for each function evaluation, and we use a high-performance computing cluster to do the job.

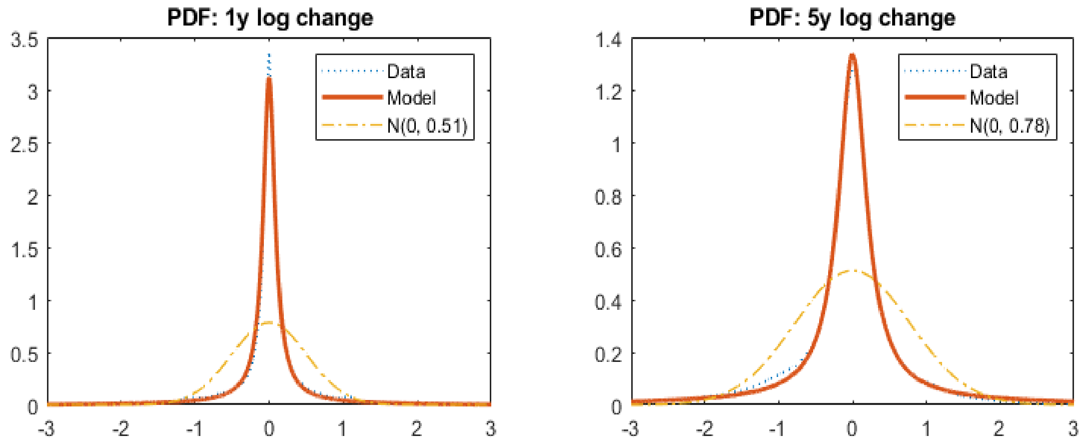
shape parameter $2\rho_\ell L/\sigma_\ell^2$ and a scale parameter $2\rho_\ell/\sigma_\ell^2$. However, we do not have an analytical solution for the stationary distribution of the changes of its logarithmic process, and thus we need to simulate the process. In each of the simulated samples, we draw from the known stationary distribution of the square-root process and simulate the process for two years. Then, we calculate the changes of log (annual) earnings as well as a few important statistics, which are averaged across samples. To reduce the variability in the SMM procedure contributed by extreme values of the Gamma distribution that are not easily sampled, we use importance sampling with exponential tilting adjustment on the shape and scale parameters.[20]

In the SMM procedure, we consider the standard deviation, the skewness, as well as the fractions of households whose log earnings changes are less than 5%, 10%, and 20%. We minimize the sum of the squared errors of these statistics with a certain weighting matrix. The weighting matrix starts from an identity matrix and is updated once (instead of searching for the optimal weighting matrix, we iterate until the minimizers converge). We find that the statistics from the simulated log earnings changes match the data reasonably well, except for the skewness. Guvenen et al (2021) argue that the negative skewness in the data could be entirely due to unemployment (disaster) shocks, not captured in our model. In Guvenen et al (2021), the negative skewness of -1.07 as measured by the third central moment may be driven by extreme observations. To take this issue into account, they compute Kelley's skewness, which is robust to observations above the 90th or below the 10th percentile of the distribution. They find that the asymmetry (negative Kelley's skewness) is prevalent across the entire distribution rather than being driven just by the tails. We find that Kelley's skewness is close to zero (-0.0001) in our simulated model. Thus the density of our model simulated log earnings changes is almost symmetric.

In Table 1, we also report the kurtosis (15.94) from our simulated model, which is close to the data

[20]For each pair of parameters of the Gamma distribution, we follow the common procedure in importance sampling to search for the tilted parameters that minimize the variance.

Figure 2: Histograms of one- and five-year log earnings changes



Notes: This figure plots the densities of one- and five-year earnings changes from the data of Guvenen et al (2021) and from our simulated model, superimposed on Gaussian densities with the same standard deviation.

(14.93). Figure 2 plots the histograms of 1-year and 5-year changes of log earnings from our simulated model and from the data of Guvenen et al (2021), superimposed on Gaussian densities with the same standard deviation. The figure shows that our model densities are close to the data.

Notice that as we normalize the equilibrium wage to 1, the (after-tax) labor income process is the same as the employment-shock process. This does not have any impact on the estimation above using changes of log earnings, since the steady-state wage rate is canceled. Our result shows that the employment shock process can also match the cross-sectional data.[21]

It merits emphasis that the square-root earnings process does not feature a Pareto tail, but it matches reasonably well many statistics in the world-wide inequality database (WID). This process itself generates a positive skewness for the earnings levels. The very top earners in the model have reasonable labor income shares compared to the WID. The average earnings of the top 0.1% people are about 115.6 times of those of the bottom 50% people, while the WID has a factor of about 150-200 (Piketty, Saez, and Zucman (2018)). Nevertheless, as will be shown later, we can generate an equilibrium wealth distribution that is much more right skewed and has a much thicker tail than the earnings distribution, due to the capital income jumps.

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 about 13%. It should be noted

[21] In a previous version, we use the limiting Gamma distribution to match the PSID data. Thanks to the comments by the Editor and one anonymous referee, the estimation using the changes of log earnings improves the quantitative results significantly.

that our model allows for both success and failure of innovations or R&D, because the jump returns may not be enough to compensate the loss arising from maintenance costs. This can happen if the actual 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 probabilities, and recalibrate parameters. We find that the model implication for the wealth distribution statistics does not change significantly, consistent with the discussion above about the analytical features of the tails.

The jump size distribution ν is important to match the wealth distribution in the data. We adopt the HED specification in (36). An important advantage of the HED is that it is analytically tractable as its moment generating function has the following closed form

$$\mathbb{E}_{\nu}[\exp(tq)] = \sum_{j=1}^{n} \frac{p_j}{1-\mu_j t}, \tag{41}$$

for $t < \min_j\{1/\mu_j\}$. Thus the Laplace transform $\mathbb{E}_{\nu}[\exp(-\alpha q)]$ exists and has a simple analytical expression for all $\alpha > 0$. This expression will be used in the household decision rules (7) and (11). Also notice that the HED in (36) 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 will show later that our calibrated model can still match reasonably well the wealth shares in the data.

We consider $n = 2$ components in the HED and choose values of μ_1, μ_2, and p_1 (note $p_2 = 1 - p_1$) to target three statistics: 13.5% of the average pre-tax private returns to innovations and/or R&D (i.e., $\lambda_k \mathbb{E}_{\nu}\left[q\right]/(1-\tau_k)$ in the model), and top 0.1% and 20% wealth shares in the US data. Griffith (2000) documents that the private return is about 27% and can range from 10% to 30% in the US. The public return can be even higher. Our target of 13.5% for the private return is conservative. Using administrative tax data, Smith, Zidar, and Zwick (2021) estimate that the top 0.1% wealth share increased from 9.9% in 1989 to 15% in 2016. 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 14.7% (the average between 2000 and 2016) as our target for the top 0.1% wealth share. We also choose the top 20% wealth share, which is 79.5%, a conventional target according to Castaneda, Diaz-Gimenez, and Rios-Rull (2003).

In our exercise, targeting the top 0.1% showcases the model's ability to match the right tail of the distribution, and targeting the top 20% illustrates the model's ability to capture the general features of inequality (e.g., the conventional view of 80-20 rule). Later, we assess the model's performance in other statistics. For example, the top 1% and 10% wealth shares are 31.8% and 66.7%, respectively. Our model turns out to match these non-targeted statistics well.

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 in (25). The time increment represents one week. In the end, we run 100 simulations of the wealth process x_t, with each simulation having 15 years (52 weeks per year) and 100,000 people. Increasing the simulation length and/or the number of people does not change our results significantly. In equation (25), important parameters that govern the wealth distribution are $\mu_x = -0.1367$, $\rho_x = 0.1750$, and $\phi = 0.7537$ according to our calibration.

Table 2: Calibrated parameter values

	Value	Explanation/Target		Value	Explanation/Target
β	0.1417	MPC = 0.20	B	1.5798	$B/Y = 0.81$
γ	4.2250	relative risk aversion 5	G	0.3706	$G/Y = 0.19$
ψ	1.5000	EIS	μ_2	409.25	top 0.1% wealth share
α	0.3300	capital share	μ_1	0.1357	top 20% wealth share
δ	0.1197	$I/Y = 0.16$	p_2	0.0046	average innovation return 13.5%
A	1.3495	$w = 1$	p_1	0.9954	$1 - p_2$
L	0.8033	estimated	η	0.0047	$R^k - r = 3.7\%$
ρ_ℓ	0.0033	estimated	χ	0.0251	interest rate $r = 2.5\%$
σ_ℓ	0.1098	estimated	λ_k	0.0500	innovation probability
$\tau_\ell = \tau_k$	0.2500	average tax rate			

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 and 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 the component distributions, the more statistics the model can match. However, as n increases, the 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 $n = 2$ to $n = 4$ and above (we experimented with $n = 3$ and the difference is already small) requires substantially more people in our simulation, which quickly becomes infeasible even for medium-scale high-performance computing facilities in a normal research institution, while the model outcome is not substantially improved.[22]

[22] The simulation has been tested in a UCL high-performance computing cluster. We optimize the procedure with parallel computing via Matlab. We use 32 i7 cores and about 200GB memories.

4.2 Results and Sensitivity Analysis

Table 3 presents the baseline quantitative steady-state results based on our calibration. Though not targeted, our model generates about 1.7% wealth share for the bottom 50%, which is the same as in the data. Our model also generates 34.8% 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, some difference from the data is expected. Notice that top 1% and top 10% wealth shares in the model are also close to the data. Notably, our model can match the right tail of the wealth distribution not at the expense of missing other parts of the distribution. In particular, our calibrated model can match the entire wealth distribution reasonably well.

As is well known in the literature, it is notoriously challenging for an equilibrium model to match the extreme right tail of the wealth distribution in the data. The baseline simulation of our model generates 14.7% 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 by not too much. The reason is that the 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.

Table 3: Wealth distribution statistics (Wealth shares and Gini)

	Top 0.1%	Top 1%	Top 10%	Top 20%	Top 50-10%	Bottom 50%	Gini
US data	14.7%	31.8%	66.7%	79.5%	32.0%	1.7%	0.824
Model	14.7%	33.5%	63.5%	79.5%	34.8%	1.7%	0.794

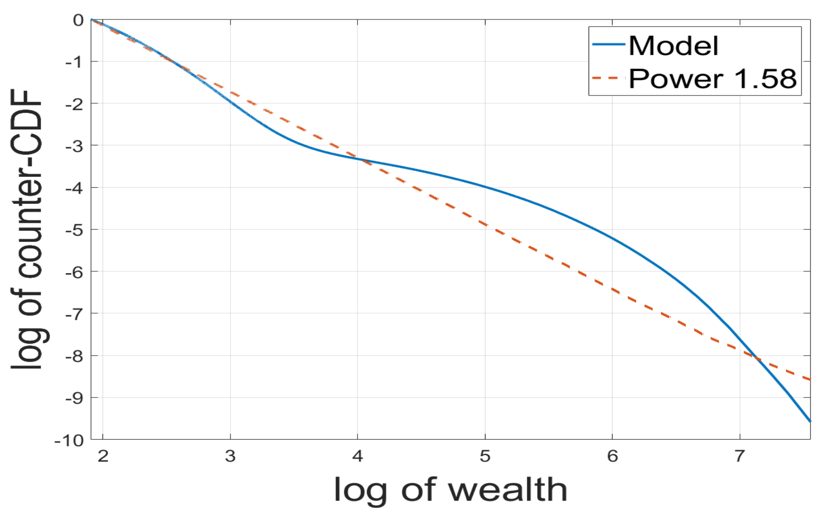
Notes: The wealth shares of top 0.1%, top 1%, and top 10% are 2000-2016 averages from Smith, Zidar, and Zwick (2021). The top 20% data is from Table 2 of Castaneda, Diaz-Gimenez, and Rios-Rull (2003) estimated from Survey of Consumer Finances. Gini coefficient is the 2000-2016 average of data from WID. The rest are the averages between 2000 and 2016 obtained from the distributional financial account of the Federal Reserve Board. The model statistics are averages of 100 simulations of results obtained with 15 model years and with dt approximated by one week, and each simulation has 100,000 people starting from the same initial levels of labor income and wealth.

Another way of looking at the tail of the wealth distribution is to plot the counter CDF in log against the wealth level in log. As is well known, the power law or the Pareto distribution of wealth implies a negative linear relationship between the two. In Figure 3, we plot this relationship using a simulated power law with the power parameter 1.58 and the lower bound estimated from the simulated wealth data in our model.[23] Notice that the estimate is close to the US data (see footnote 4). Without assuming a

[23] We use the Matlab code plfit.m obtained from http://www.santafe.edu/ aaronc/powerlaws/. This routine estimates the power parameter and the lower bound according to the goodness-of-fit based method described in Clauset, Shalizi, and

power law, we also plot the log-log relationship using the same simulated wealth data, which are known to be generated from an exponential tailed distribution by Proposition 6. The result suggests that the exponential tailed distribution of wealth generated from the model is indeed close to a Pareto distribution in finite samples, with sometimes undershooting and other times overshooting. When we increase the number of exponential distributions in the HED for the jump size, the model generated log-log line is even closer to that of a Pareto distribution.

Figure 3: Wealth distribution tail



Notes: This figure plots the log of the counter CDF of wealth against the wealth level in log. Specifying a power law distribution for the model simulated wealth data, we estimate the power parameter to be 1.58. The downward straight line corresponds to this power law. The curve corresponds to the empirical distribution derived from the same simulated wealth data.

Our model generated wealth Gini coefficient is 0.794, slightly smaller than the data (0.824). Some households are in debt in our simulations and these households do not receive government transfers. There are roughly 24% of people having negative wealth, although none of them have negative consumption.

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, and the Poisson arrival rate λ_k (see Tables 4 and 5). When changing one parameter value, we keep other parameter values fixed as in the benchmark calibration in Table 2. The sensitivity analysis helps us better understand our model mechanics.

We first examine the impact of the EIS ψ. Intuitively, consumption/saving behavior depends on the EIS. On the one hand, the higher the EIS ψ, the larger the substitution effect and thus the lower the saving under incomplete markets with $r < \beta$. As a result, the equilibrium interest rate r rises with the

Newman (2009).

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), because $\theta = [\psi(\beta - r) + r]^{\frac{1}{1-\psi}}$ in (9) depends on ψ. The value of θ is very sensitive to a small change of ψ under our calibration compared to changes of other parameter values. 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 saving 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.

Table 4: Sensitivity analysis for EIS and risk aversion

	Capital	Wealth	r (%)	MPC (%)	Bottom 50% share (%)	Top 10% share (%)	Top 1% share (%)	Top 0.1% share (%)	Gini
Benchmark	2.61	4.18	2.50	20.00	1.7	63.5	33.5	14.7	0.794
ψ =1.60	2.51	4.09	2.98	20.88	6.7	59.1	32.4	14.7	0.717
– and PE	2.61	3.34	2.50	21.17	-1.9	69.4	39.4	18.0	0.862
ψ =1.70	2.40	3.98	3.53	21.61	11.1	55.3	31.5	14.7	0.651
– and PE	2.62	2.57	2.50	22.30	-6.9	77.4	47.4	22.2	0.957
γ =3	2.58	4.17	2.63	19.94	3.1	62.5	33.7	14.9	0.773
– and PE	2.61	3.94	2.50	20.00	0.8	65.2	35.3	15.6	0.812
γ =2	2.54	4.12	2.82	19.84	5.2	60.9	33.6	15.0	0.742
– and PE	2.62	3.60	2.50	20.00	-0.7	67.8	38.3	17.1	0.841

Notes: For each parameter, the corresponding row shows the result in the stationary equilibrium; the row labeled "- and PE" shows results of the corresponding economy with the prices r and w fixed at the levels of the benchmark economy.

What is the impact on the MPC $\vartheta = \psi(\beta - r) + r$ in our model? An increase in ψ raises the MPC for a fixed $r \in (0, \beta)$.[24] As r rises in general equilibrium, this effect reduces ϑ given $\psi > 1$. In our numerical experiment, the former direct effect dominates. Therefore, the MPC increases with ψ in general equilibrium. This result also explains why aggregate saving decreases with ψ as discussed before.

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 the capital stock K, causing the capital income jump intensity $\lambda = \lambda_k K$ to decline, and thereby reducing the top wealth shares and wealth inequality. Second, a decrease in the capital stock raises the marginal product of capital R^k and hence raises wealth inequality.

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

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, when 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 from the four channels is ambiguous. Table 4 shows that the top 10%, 1%, and 0.1% wealth shares as well as the wealth Gini coefficient all decline (some to the third decimal place) with ψ in general equilibrium, but the bottom 50% wealth share increases with ψ. Therefore, the net effect is to reduce wealth inequality.

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 and the last three channels vanish. By (7) or (12), capital demand increases with ψ as the jump risk premium decreases with ψ. Working in the opposite direction of the first channel generates more wealth inequality due to higher aggregate capital in partial equilibrium. Moreover, the bottom 50% people are in debt; for example, their debt is 1.9% of aggregate wealth for $\psi = 1.60$. This is in contrast to the case in general equilibrium, in which the increased interest rate (generated by the increased ψ) induces the poor people to save.

Next we consider the impact of the risk aversion parameter γ. Table 4 shows that the impact of a 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 $\psi > 1$. The MPC thus falls when we reduce γ. Therefore, the relationship between inequality and MPC depends on the source generating the MPC variation. Unlike in the case for the EIS ψ, the aggregate and distributional effects are much less sensitive to changes in γ.

Table 5: Sensitivity analysis for jump intensity

	Capital	Wealth	r (%)	MPC (%)	Bottom 50% (%)	Top 10% (%)	Top 1% (%)	Top 0.1% (%)	Gini
Benchmark	2.61	4.18	2.50	20.00	1.7	63.5	33.5	14.7	0.794
λ_k =0.055	2.59	4.16	2.60	19.95	-0.3	66.5	36.2	15.4	0.829
– and PE	2.61	3.99	2.50	20.00	-2.4	69.1	38.0	16.2	0.866
λ_k =0.060	2.57	4.15	2.68	19.91	-2.1	69.5	39.0	16.0	0.865
– and PE	2.61	3.77	2.50	20.00	-7.0	75.1	42.6	17.5	0.945

Notes: For each parameter, the corresponding row shows the result in the stationary equilibrium; the row labeled "- and PE" shows results of the corresponding economy with the prices r and w fixed at the levels of the benchmark economy.

Finally, we examine the role of jump risks by focusing on the jump intensity parameter λ_k in Table 5. In partial equilibrium with fixed prices, a small increase in λ_k induces households to raise capital demand very slightly, but reduce bond holdings, thereby raising wealth inequality. This effect can also be seen from the negative wealth share of the bottom 50% and the higher wealth share at the very top 0.1% due to more frequent capital income jumps.

In general equilibrium, a higher λ_k pushes the interest rate above the baseline level because of the decline in savings as well as a slightly higher capital demand. As a result, the equilibrium level of capital falls slightly. Facing a higher interest rate, aggregate wealth is higher and the wealth inequality is less severe, compared to the case of partial equilibrium. Consistent with the effect of a higher interest rate on the saving behavior, the MPC slightly falls.

We also experiment with increasing μ_2, the larger mean of the component of the jump size distribution, by 5% and 10%. The qualitative effects on the aggregate and on the distribution are the same as increasing λ_k. For simplicity, we do not report this result in Table 5.

5 Conclusion

In this paper we have provided a tractable heterogeneous-agent model with incomplete markets in continuous time; the model can generate a heavier wealth tail than that of the labor income distribution. We can explain the wealth inequality observed in the data even if our model generates a wealth distribution that has an exponential tail. Because Pareto-tailed and some exponential-tailed distributions are almost indistinguishable in the data with a finite sample, our calibrated model can match the wealth distribution in the data reasonably well. Our key story is that rich people can build wealth from rare capital 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. Our theory provides a new source of wealth inequality and has implications for future empirical research. It would also be interesting for future research to apply our theory to study how tax policy can affect the wealth inequality.[25]

[25] A previous version of the paper studies the aggregate and distributional impacts of various capital tax, which is available upon request.

Appendix

A 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 point process with intensity $\lambda_t = \lambda_k k_t$. By Appendix B, the HJB equation is given by

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

Conjecture that the value function takes the form:

$$V_t = V\left(x_t, \ell_t\right) = \theta\left(x_t + \xi_\ell \ell_t + \xi_0\right), \quad \text{(A.2)}$$

where θ, ξ_ℓ, and ξ_0 are constants to be determined. By Ito's Lemma, it follows from (B.3) and (B.4) that

$$\mu_t = \theta r x_t + \theta\left(R^k - \chi - r\right)k_t - \frac{\theta\eta}{2}k_t^2 - \theta c_t + \theta\Upsilon_t + \theta w \ell_t + \theta\xi_\ell\rho_\ell\left(L - \ell_t\right), \quad \text{(A.3)}$$

$$\sigma_t^W = \theta\xi_\ell\sigma_\ell\sqrt{\ell_t}, \quad \sigma_t^J = \theta q_t. \quad \text{(A.4)}$$

Plugging the above equations into (A.1) and taking first-order conditions, we have

$$f'\left(c_t\right) = \theta f'\left(V_t\right), \quad \text{(A.5)}$$

$$\theta\left(R^k - \chi - r\right) - \theta\eta k_t + \frac{\lambda_k \mathbb{E}_\nu\left(u\left(V_t + \sigma_t^J\right) - u\left(V_t\right)\right)}{u'\left(V_t\right)} = 0. \quad \text{(A.6)}$$

By (2) and (A.2), we have

$$c_t = \theta^{-\psi}V_t = \theta^{1-\psi}\left(x_t + \xi_\ell\ell_t + \xi_0\right), \quad \text{(A.7)}$$

$$k_t = \frac{R^k - \chi - r}{\eta} + \frac{\lambda_k\mathbb{E}_\nu\left[1 - \exp\left(-\gamma\theta q\right)\right]}{\eta\gamma\theta}. \quad \text{(A.8)}$$

We then obtain (6) and (7).

We now verify the conjecture (A.2) and derive coefficients. Using (2) and (A.7), we divide the two sides of (A.1) 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 (A.3), (A.4), (A.2), and (A.7) into the preceding equation, we obtain

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

Matching coefficients yields

$$\frac{\left(\beta-\theta^{1-\psi}\right)}{1-1/\psi}=r-\theta^{1-\psi},\tag{A.9}$$

$$\frac{\left(\beta-\theta^{1-\psi}\right)}{1-1/\psi}\xi_\ell=w-\left[\rho_\ell+\theta^{1-\psi}\right]\xi_\ell-\frac{\gamma}{2}\theta\left(\xi_\ell\sigma_\ell\right)^2,\tag{A.10}$$

$$\begin{aligned}\frac{\left(\beta-\theta^{1-\psi}\right)}{1-1/\psi}\xi_0 \;=\;& \left(R^k-r-\chi\right)k_t-\frac{\eta}{2}k_t^2-\theta^{1-\psi}\xi_0+\Upsilon_t\\ &+\xi_\ell\rho_\ell L-\frac{\lambda_t}{\gamma\theta}\mathbb{E}_\nu\left[\exp\left(-\gamma\theta q\right)-1\right].\end{aligned}\tag{A.11}$$

Equation (A.9) gives (9). Simplifying equation (A.10) 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 (A.11) gives (11).

To show $a_h=\frac{(r+\rho_\ell)\xi_\ell}{w}\in(0,1)$, define the function

$$g\left(x\right)\equiv\frac{\gamma\theta}{2}\left(x\sigma_\ell\right)^2+\left(r+\rho_\ell\right)x-w.$$

Since $g\left(0\right)=-w<0$ and $g\left(\frac{w}{r+\rho_\ell}\right)=\frac{\gamma\theta}{2}\left(\frac{w\sigma_\ell}{r+\rho_\ell}\right)^2>0$, it follows from the intermediate value theorem that the positive root ξ_ℓ satisfies $0<\xi_\ell<\frac{w}{r+\rho_\ell}$, which is equivalent to $a_h=\frac{(r+\rho_\ell)\xi_\ell}{w}\in(0,1)$.

Finally, we need to establish the transversality condition. Unfortunately, it is technically challenging to establish this condition for recursive utility. We leave this for future research and refer the reader to Fleming and Soner (2006) and Beare, Seo, and Toda (2022) for a rigorous treatment in the case of standard utility functions. Q.E.D.

Proof of Proposition 2: Optimal wealth x_t satisfies (25). Solving this equation yields

$$x_t = x_0 e^{-\rho_x t} + \mu_x \frac{1-e^{-\rho_x t}}{\rho_x} + \int_0^t e^{-\rho_x (t-s)} dJ_s + \int_0^t e^{-\rho_x (t-s)} \phi w \ell_s ds.$$

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 as $t \to \infty$ 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

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

By the Binomial expansion formula

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

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

$$\begin{aligned}
\mathbb{E}\left[\widetilde{x}_t^m \widetilde{z}_t^n\right] &= e^{-\kappa_{m,n}t}\mathbb{E}\left[\widetilde{x}_0^m \widetilde{z}_0^n\right] + \int_0^t e^{-\kappa_{m,n}(t-s)} Q(m,n)\, ds \\
&= e^{-\kappa_{m,n}t}\mathbb{E}\left[\widetilde{x}_0^m \widetilde{z}_0^n\right] + \frac{1}{\kappa_{m,n}}\left(1 - e^{-\kappa_{m,n}t}\right) Q(m,n),
\end{aligned} \tag{A.12}$$

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

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

Since $\kappa_{m,n} > 0$, taking limits in (A.12) 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 (31).

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 $\zeta_j = \mathbb{E}_\nu [q^j]$ exists for $1 \le j \le j^*$. Whenever $\zeta_j = \mathbb{E}_\nu [q^j]$ 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:

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

The labor income skewness and (excess) kurtosis are given by

$$
\begin{aligned}
\text{Skew}[z] &= \frac{M_{0,3}}{(M_{0,2})^{3/2}} = \frac{\sigma_z^2}{\rho_\ell} (M_{0,2})^{-1/2} = \left(\frac{2\sigma_z^2}{\rho_\ell Z} \right)^{\frac{1}{2}}, \\
\text{Kurt}[z] &= \frac{M_{0,4}}{(M_{0,2})^2} - 3 = \frac{3\sigma_z^2}{\rho_\ell Z}.
\end{aligned}
$$

Step 2. Applying Proposition 3, we can derive

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

Since

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

we have

$$\frac{\mathrm{Var}[x]}{\mathrm{Var}[z]} = \frac{M_{2,0}}{M_{0,2}} = \frac{\phi^2}{\rho_x(\rho_x+\rho_\ell)} + \frac{(\lambda_k K)\,\zeta_2}{2\rho_x M_{0,2}} = \frac{\phi^2}{\rho_x(\rho_x+\rho_\ell)} + \varpi_1, \tag{A.13}$$

where

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

The correlation between x_t and z_t is given by

$$\frac{M_{1,1}}{\sqrt{\mathrm{Var}[x]}\sqrt{\mathrm{Var}[z]}} = \frac{\phi}{\rho_x+\rho_\ell}\frac{\sqrt{\mathrm{Var}[z]}}{\sqrt{\mathrm{Var}[x]}}.$$

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

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

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

The wealth skewness is given by

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

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

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

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

Plugging the above expression for $M_{1,3}$ into above equation for $M_{2,2}$ and using Step 1 and (A.13), we can derive

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

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

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

where we have used (A.13) to derive 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 kurtosis

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

Using (A.13) , we have

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

By Step 1, we obtain the wealth kurtosis:

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

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

Proof of Proposition 6: We have the following affine jump-diffusion process

$$dx_t=-\rho_x x_t dt+\mu_x dt+\phi z_t dt+dJ_t, \tag{A.14}$$

$$dz_t=\rho_\ell\left(Z-z_t\right)dt+\sigma_z\sqrt{z_t}dW_t^\ell, \tag{A.15}$$

where μ_x is given by (26).

We compute the exponential moment

$$\mathbb{E}\left[\exp\left(u^x x_t+u^z z_t\right)\mid\left(x_0,z_0\right)=\left(x,z\right)\right]=\exp\left(A\left(t\right)+B^x\left(t\right)x+B^z\left(t\right)z\right), \tag{A.16}$$

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

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

$$\dot{A}\left(t\right)=B^x\left(t\right)\mu_x+\rho_\ell Z B^z\left(t\right)+\lambda_k k\mathbb{E}_\nu\left[\exp\left(B^x\left(t\right)q\right)-1\right], \tag{A.17}$$

$$\dot{B}^x\left(t\right)=-\rho_x B^x\left(t\right), \tag{A.18}$$

$$\dot{B}^z\left(t\right)=-\rho_\ell B^z\left(t\right)+B^x\left(t\right)\phi+\frac{1}{2}\left(B^z\left(t\right)\sigma_z\right)^2. \tag{A.19}$$

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

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

There are two equilibrium points of the ODE system (A.18) and (A.19) for $\left(B^z\left(t\right),B^x\left(t\right)\right):\left(0,0\right)$ and $\left(2\rho_\ell/\sigma_z^2,0\right)$ (see Figure 4). 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 $\mathbb{R}_+\times\left(-\infty,\min_j\left\{1/\mu_j\right\}\right)$ gives the stable set for the system (A.17)-(A.19).

By Keller–Ressell and Mayerhofer (2015), the set of (u^x, u^z) such that

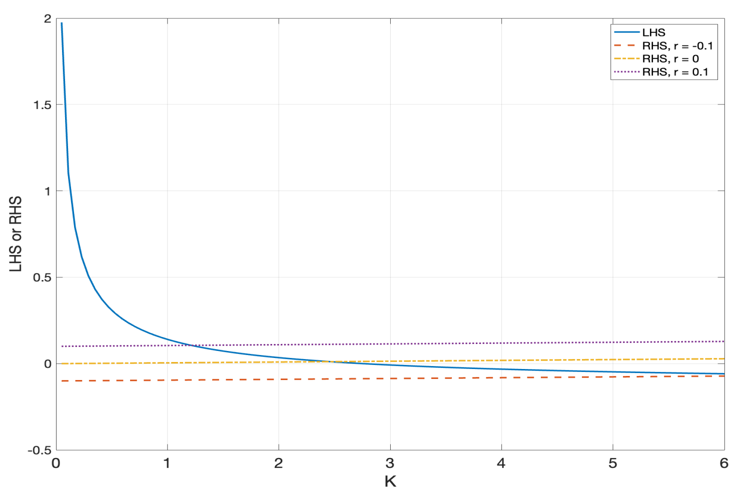
$$\lim_{t\to\infty} \mathbb{E}\left[\exp\left(u^x x_t + u^z z_t\right) \mid (x_0, z_0) = (x, z)\right] < \infty$$

is the same as the stable set of the system of ODEs (A.17), (A.18), and (A.19). Since this set contains a neighborhood of the origin,

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

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 4: Vector fields



Proof of Proposition 7: We continue the analysis in the proof of Proposition 6 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 point. 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 4). 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 4. Then for any $B^z(0) = u^z$, there exists a unique initial value $B^x(0) = u^x = g\left(u^z\right)$ such that the ODE system (A.18) and (A.19) has a unique saddle-path solution for $B^z(t)$ and $B^x(t) = g\left(B^z(t)\right)$ that converges to the equilibrium point $(2\rho_\ell/\sigma_z^2, 0)$.

The stable set S_0 for the ODE system for $\left(B^z(t), B^x(t)\right) \in \mathbb{R}_+^2$ is given by the region in the figure, whose nonlinear boundary is the saddle path. Let S denote the intersection of S_0 and $\mathbb{R}_+ \times (-\infty, \min_j \{1/\mu_j\})$. Note that for $A(t)$ in (A.17) to converge, we must have $B^x(0) = u^x < \min_j \{1/\mu_j\}$.

Then S is the stable set for the ODE system (A.17), (A.18), and (A.19). For any $(B^z(0), B^x(0))$ in S, we have

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

Moreover,

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

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

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

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

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

for any $\alpha(0,1) \in S$ and $\alpha > 0$. Then we can compute the maximum α such that $\alpha(0,1) \in S$, 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 $g(0)$, we have

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

Now consider the labor income process z_t. We have

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

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

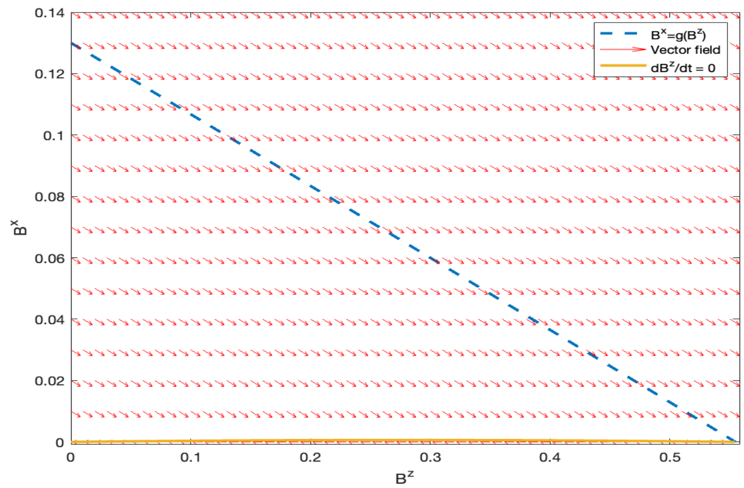
$$-\rho_\ell \alpha + \frac{1}{2}(\alpha\sigma_z)^2 = 0.$$

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

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

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 5 shows one numerical example. Since θ increases with r and $\mathbb{E}_\nu[1 - \exp(-\gamma\theta q)]/\theta$ decreases with θ, we have $\mathbb{E}_\nu[1 - \exp(-\gamma\theta q)]/\theta$ decreases with r. Thus the line, denoted by RHS in Figure 5, shifts up as r increases. Thus $K(r)$ decreases with r.

Figure 5: 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 (23) into (38) yields

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

Using equations (17) and (20), we can derive

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

where the last equality follows from (7) and $k = K$. After using the government budget constraint, we have

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

Notice that the term $rB + \Upsilon$ reflects the Ricardian equivalence in the aggregate discussed before.

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 $(0, 1)$. If G is sufficiently small, then $rB + \Upsilon > 0$. 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 8: By Lemmas 1 and 2 and the intermediate value theorem, there exists a solution $r \in (0, \beta)$ to equation (40). Q.E.D.

B 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:

$$U_t = f^{-1} [f(c_t) dt + \exp(-\beta dt) f(\mathcal{R}_t(U_{t+dt}))], \quad \text{(B.1)}$$

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

$$dU_t = \mu_t dt + \sigma_t^W dW_t + \sigma_t^J dN_t, \quad \text{(B.2)}$$

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 (σ_t^W, σ_t^J) 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

$$\begin{aligned}\mu_t &= U_x(x_t,\ell_t)\left[rx_t+\left(R^k-r-\chi-\frac{\eta}{2}k_t\right)k_t+\Upsilon+w\ell_t-c_t\right] \\ &\quad +U_\ell(x_t,\ell_t)\rho_\ell(L-\ell_t)+\frac{1}{2}U_{\ell\ell}(x_t,\ell_t)\sigma_\ell^2\ell_t, \end{aligned} \tag{B.3}$$

$$\sigma_t^W = U_\ell(x_t,\ell_t)\sigma_\ell\sqrt{\ell_t},\ \ \sigma_t^J = U(x_t+q_t,\ell_t)-U(x_t,\ell_t). \tag{B.4}$$

Given a small time increment dt, we heuristically write $dU_t = U_{t+dt} - U_t$. Then we rewrite (B.2) 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

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

Taking conditional expectations yields

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

where $\lambda_t = \lambda_k k_t$.

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

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

Following the same procedure again gives

$$f(\mathcal{R}_t(U_{t+dt})) = f(U_t)+f'(U_t)\left[\mu_t+\frac{1}{2}\frac{u''(U_t)}{u'(U_t)}\sigma_t^W\left(\sigma_t^W\right)'+\lambda_t\frac{\mathbb{E}_\nu\left(u\left(U_t+\sigma_t^J\right)-u(U_t)\right)}{u'(U_t)}\right]dt. \tag{B.5}$$

Up to a first-order approximation, we have

$$\exp(-\beta dt) = 1-\beta dt. \tag{B.6}$$

Plugging (B.5) and (B.6) into (B.1), we can derive

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

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

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

C Pareto versus Exponential Tails

In this appendix we provide formal definitions of tail thickness and discuss the literature on the Pareto tails. We then discuss the difference between the Pareto and exponential tails.

Definitions

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. 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

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

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.

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

$$dx_t = R_t^x x_t dt + y_t dt - c_t dt, \tag{C.1}$$

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 (C.1) becomes

$$dx_t = (r - \vartheta)\, x_t dt + (y_t - \Phi_t)\, dt. \tag{C.2}$$

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 $(y_t - \Phi_t)$. Stachurski and Toda (2019) prove that under some standard BHA assumptions, the wealth distribution inherits the tail behavior of the labor income process. (See also Grey (1994), Benhabib, Bisin, and Luo (2017) and Benhabib and Bisin (2018)). 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. Notice that a geometric Brownian motion x_t stopped at an exponentially distributed time has a Pareto tail, but $\ln x_t$ has an exponential tail. Based on this observation, Beare et al (2022) study the tail probabilities of a light-tailed Markov-modulated Levy process stopped at a state-dependent Poisson rate. They show that the tails decay exponentially at rates given by the unique positive and negative roots of the spectral abscissa of a certain matrix-valued function.

For both types of processes above, 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 (2000) in the finance literature. We introduce capital income jumps that appear in equation (25) 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 (C.2) 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). His results confirm the finding of Stachurski and Toda (2019).

References

Achdou, Yves, Jiequn Han, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, 2022. Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. Review of Economic Studies, 89(1), pp. 45–86.

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

Angeletos, George-Marios, 2007. Uninsured Idiosyncratic Investment Risk and Aggregate Saving. Review of Economic Dynamics, 10(1), pp.1-30.

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

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

Arellano, Manuel, Stéphane Bonhomme, Micole De Vera, Laura Hospido, and Siqi Wei, 2022. Income risk inequality: Evidence from Spanish administrative records. Quantitative Economics, 13(4), pp. 1747-1801.

Beare, Brendan K., Won-Ki Seo, and Alexis Akira Toda, 2022. Tail Behavior of Stopped Lévy Processes with Markov Modulation. Econometric Theory, 38(5), pp. 986–1013.

Benhabib, Jess and Alberto Bisin, 2018. Skewed Wealth Distributions: Theory and Empirics. Journal of Economic Literature, 56, pp. 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, pp. 123-157.

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

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

Benhabib, Jess, Alberto Bisin, and Mi Luo, 2017. Earnings Inequality and Other Determinants of Wealth Inequality. American Economic Review Papers & Proceedings, 107, pp. 593-597.

– 2019. Wealth Distribution and Social Mobility in the US: A Quantitative Approach. American Economic Review, 109, pp. 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, pp. 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, pp. 261-331.

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

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

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

Caldara, Dario, Jesus Fernadez-Villaverde, Juan F. Rubio-Ramrez, and Wen Yao, 2012. Computing DSGE models with recursive preferences and stochastic volatility. Review of Economic Dynamics, 15, pp. 188-206.

Calvet, Laurent-Emmanuel, 2001. Incomplete markets and volatility. Journal of Economic Theory, 98(2), pp. 295-338.

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

Carroll, Christopher D. and Miles S. Kimball, 1996. On the Concavity of the Consumption Function. Econometrica, 64(4), pp. 981-992.

Carroll, Christopher D., Jiri Slacalek, and Kiichi Tokuoka, 2014. The Distribution of Wealth and the MPC: Implications of New European Data. American Economic Review, 104(5), pp. 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, pp. 818-857.

Clauset, Aaron, Cosma Rohilla Shalizi, and M. E. J. Newman, 2009. Power-Law Distributions in Empirical Data. SIAM Review, 51(4), pp. 661-703.

Clementi, F. and M. Gallegati, 2005. Power Law Tails in the Italian Personal Income Distribution. Physica A: Statistical Mechanics and Theoretical Physics, 350, pp. 427-438.

Cowell, Frank A. 2011. Inequality among the Wealthy. Centre for Analysis of Social Exclusion, LSE CASE Papers, 150.

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

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

Dagsvik, John, Jia, Zhiyang, Vatne, Bjorn and Weizhen Zhu, 2013. Is the Pareto–Lévy law a good representation of income distributions? Empirical Economics, pp. 719–737.

De Nardi, Mariacristina, 2004. Wealth Inequality and Intergenerational Links. Review of Economic Studies, 71, pp. 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, pp. 824-57.

De Vries, Tjeerd and Alexis Akira Toda, 2021. Capital and labor income Pareto exponents across time and space. Review of Income and Wealth, 8(4), pp.1058-1078.

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

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

Epstein, Larry G. and Stanley E. Zin, 1989. Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework. Econometrica, 57, pp. 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, pp. 255-294.

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

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

Gouin-Bonenfant, Emilien and Alexis Akira Toda, 2023. Pareto extrapolation: An analytical framework for studying tail inequality. Quantitative Economics, 14, pp. 201-233.

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, pp. 169-83.

Guvenen, Fatih, 2011. Macroeconomics with Heterogeneity: A Practical Guide. Federal Reserve Bank of Richmond's Economic Quarterly, 97, pp. 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? Econometrica, 5, pp. 2303-2339.

Griffith, Rachel, 2000. How important is business R&D for economic growth and should the government subsidise it? Working Paper, Institute for Fiscal Studies.

Hall, Robert, 1978. Stochastic Implications of the Life Cycle-Permanent Income Hypothesis: Theory and Evidence. Journal of Political Economy, 86, pp. 971-987.

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

Heyde, C.C. and S. G. Kou, 2004. On the Controversy over Tailweight of Distributions. Operations Research Letters, 32, pp. 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, pp. 391-455.

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

Halvorsen, Elin., Joachim Hubmer, Serdar Ozkan, and Sergio Salgado, 2023. Why Are the Wealthiest So Wealthy? A Longitudinal Empirical Investigation. St Louis Fed Working Paper 2023-004A.

Jena, Rudra P., Kyoung-Kuk Kim, and Hao Xing, 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, 2015. Pareto and Piketty: The Macroeconomics of Top Income and Wealth Inequality. Journal of Economic Perspectives, 29, pp. 29-46.

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

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

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

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

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

Kesten, Harry, 1973. Random Difference Equations and Renewal Theory for Products of Random Matrices. Acta Mathematica, 131, pp. 207-248.

Klass, Oren, Biham, Ofer, Levy, Moshe, Levy, Malcai, Ofer, and Sorin Solomon, 2006. The Forbes 400 and the Pareto Wealth Distribution. Economics Letters 90, pp. 290-295.

Krishnamurthy, A. and A. Vissing-Jorgensen 2012. The Aggregate Demand for Treasury Debt. Journal of Political Economy, 120, pp. 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, pp. 842-921.

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

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

Moll, Benjamin, Lukasz Rachel, and Pascual Restrepo, 2022. Uneven Growth: Automations Impact on Income and Wealth Inequality. Econometrica, 90, pp. 2645-2683.

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

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, 2018. Distributional National Accounts: Methods and Estimates for the United States. Quarterly Journal of Economics, 133 (2), pp. 553-609.

Quadrini, Vincenzo, 2000, Entrepreneurship, Saving and Social Mobility. Review of Economic Dynamics, 3, pp. 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, pp. 1229-1302.

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

Sargent, Thomas, Neng Wang, and Jinqiang Yang, 2021. Stochastic Earnings Growth and Equilibrium Wealth Distributions, Working Paper 28473, National Bureau of Economic Research.

Smith, Matthew, Owen Zidar, and Eric Zwick, 2021. Top Wealth in America: New Estimates and Implications for Taxing the Rich. Working Paper 29374, National Bureau of Economic Research.

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), pp. 1-24.

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

Toda, Alexis Akira, 2019. Wealth distribution with random discount factors. Journal of Monetary Economics, 104, pp. 101-113.

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

Vermeulen, Philip, 2015. How Fat Is the Top Tail of the Wealth Distribution? Working Paper 1692, European Central Bank.

Wang, Neng, 2003. Caballero meets Bewley: The permanent-income hypothesis in general equilibrium. American Economic Review, 93 (3), pp. 927-936.

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

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