On the Mechanics of Top Wealth Inequality*

Fatih Guvenen Sergio Ocampo Serdar Ozkan

March 20, 2025

Abstract

Top-end wealth inequality has been receiving increasing interest in academic research and policy debates. Despite this widespread attention, there is no consensus in the academic literature about the right framework—or model ingredients—that can generate the massive concentration of wealth at the top end. In this paper, we run a *horse race* between three frameworks that are commonly used for modeling top-end inequality: (i) Infinite-horizon Aiygari-style models with "awesome state" shocks and stochastic aging, (ii) Lifecycle Aiyagari-style models with empirically estimated non-linear and non-Gaussian income processes, (iii) Lifecycle Power-law models with rate of return heterogeneity. We show several results. First, the awesome-state models with stochastic aging calibrated to match the top 1% wealth share have counterfactual implications along important dimensions: the centenarians comprise 1/4 of the population and own 2/5 of the top 1% wealth; the model can generate neither a Pareto tail nor any "hundred-millionaires" or above, among others. The second model, which features an income process matching many features of US income dynamics (taken from Guvenen, Karahan, Ozkan, and Song, 2021b), generates only 1/3 of the top 1% wealth share and essentially nobody with more than $20 million in wealth. By contrast, the return heterogeneity framework generates a thick Pareto as in the US data matches and matches other top wealth statistics while also delivering an empirically plausible rate-of-return distribution. In addition, the model reproduces the extremely rapid life cycle wealth growth at the top, matching the very high share of self-made billionaires in the data (>50%) whose wealth grows by more than 10,000-fold over their life cycle.

JEL Codes: E24, J24, J31.
Keywords: wealth inequality, Pareto tail, Power law models, awesome state shocks, income risk, non-Gaussian shocks, incomplete markets models.

*Guvenen: University of Minnesota, FRB of Minneapolis, and NBER; guvenen@umn.edu; Ocampo: University of Western Ontario; socampod@uwo.ca; Ozkan: St Louis Fed; serdar.ozkan@utoronto.ca.

1 Introduction

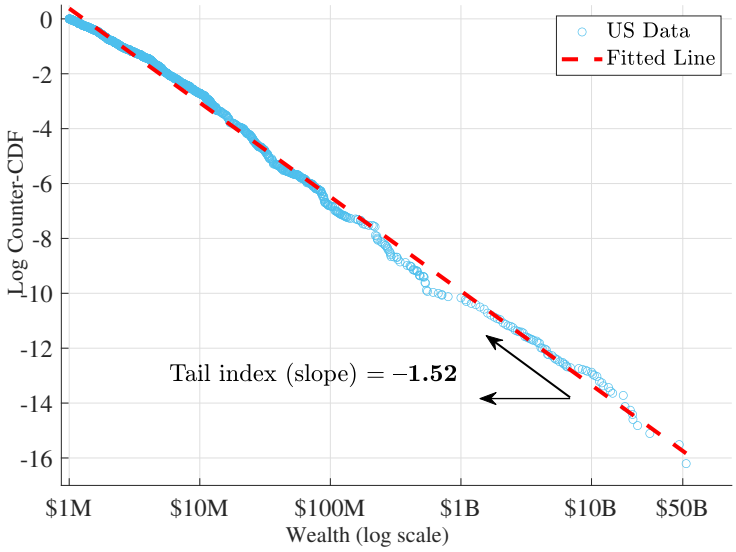
In most modern economies, wealth is extremely concentrated among a small group of households while the rest own little wealth. The right tail of the wealth distribution has a Pareto distribution with a very thick tail — implying a much higher number of super-wealthy than would be implied by lognormal distribution as Figure 1 shows. This goes beyond a statistical curiosity. Without the thickness of the wealth distribution there are no super rich. Even in the presence of high wealth concentration (e.g., high top 1% wealth shares). Furthermore, many of these super-wealthy are self-made, growing up in middle or low-income families and becoming multimillionaires or billionaires in just a few decades. Perhaps surprisingly, all these features are observed even in countries that implement strongly redistributive policies that significantly lower inequality in labor earnings.[1] Although many mechanisms have been proposed and studied, many features of wealth inequality, including those just mentioned, have proved elusive to explain.

In this paper, we systematically compare three main frameworks of heterogeneous-agent models, each featuring different subsets from half a dozen or so mechanisms proposed in the literature. We evaluate these frameworks in terms of their ability to capture the most salient features of top wealth inequality, including the Pareto tail, its thickness, the fraction who are self-made billionaires (in case they generate any billionaires), the distribution of wealth over the life cycle, among others. Further, we discuss and compare the mechanics of wealth accumulation in each framework that lead to top-end wealth concentration.

The first framework is based on the standard incomplete-markets model with idiosyncratic earnings shocks in which precautionary savings is the key motive for wealth accumulation (e.g., Aiyagari, 1994). It modifies this framework by introducing *"awesome-state" shocks* as in Castañeda, Díaz Giménez, and Ríos Rull (2003), which increase earnings by 500 to 1000 times (hence, "awesome") but are realized with a very low probability. Further, the effect of this shock lasts for a short period (5 to 10 periods on average) after which earnings fall back to their normal values. This shock is treated as a free parameter and calibrated to match statistics on wealth concentration. Models of this framework also uses either an infinite-horizon or perpetual-youth demographic

[1] For instance, Vermeulen (2016) reports that the Pareto index in the US is 1.52, higher than in Germany (1.39) and lower than in the UK (1.75) and Finland (1.88), but it is always below 2 for most developed countries. In fact, the variance of wealth (technically) does not exist in any of these countries.

FIGURE 1 – Thickness of the right tail of the US wealth distribution



Notes: Data comes from Vermeulen (2016). Tail indices are estimated from country level surveys merged with Forbes' billionaires list.

structure (i.e., agents age and die stochastically with constant probabilities). Examples of this framework include Kaymak and Poschke (2016), Greenwald, Leombroni, Lustig, and Van Nieuwerburgh (2021), Kaymak, Leung, and Poschke (2022), Boar and Midrigan (2022), and Kindermann and Krueger (2022).

The second framework considers a more empirically grounded version of the standard incomplete markets model, with a lifecycle demographic structure (which allows studying self-made wealthy) and a rich earnings process recently estimated from administrative data on earnings dynamics (from Guvenen, Karahan, Ozkan, and Song, 2021a). The process features over 20 parameters and generates *nonlinear and non-Gaussian earnings dynamics* matching 2000 empirical moments. The purpose of this second framework is to illustrate how far one can go toward explaining the features of the wealth distribution in incomplete-market models with idiosyncratic earnings shocks while remaining empirically grounded. The substantial deviations from Gaussian earnings shocks (i.e., strong negative skewness and high kurtosis) matter for wealth accumulation decisions that seek to insure against idiosyncratic income risk.[2]

[2]Guvenen, Karahan, Ozkan, and Song (2021a) find the distribution of earnings shocks displays substantial deviations from Gaussian (normal) distribution—the standard assumption in the incomplete markets literature. Negative skewness implies that workers are more likely to experience very large negative changes in their earnings (disaster shocks) compared to very large positive shocks. The high

The third framework has as its key ingredient *heterogeneity in rates of return* across households. The large magnitude and persistence of heterogeneity in returns are well documented by recent studies that use administrative data from several countries (e.g., Fagereng, Guiso, Malacrino, and Pistaferri, 2020; Bach, Calvet, and Sodini, 2020; Smith, Zidar, and Zwick, 2022), with entrepreneurial returns playing a central role in the wealth growth of those how make it and remain at the top of the distribution (Ozkan, Hubmer, Salgado, and Halvorsen, 2023). Moreover, it is well-known that heterogeneous returns generate a power law model, implying a Pareto tail in wealth (e.g., Gabaix, 2009).

Models within this framework often generate return heterogeneity from the activities of entrepreneurs subject to financial constraints (e.g., Quadrini, 2000; Cagetti and De Nardi, 2006; Cagetti and De Nardi, 2009). The recent work of Guvenen, Kambourov, Kuruscu, Ocampo, and Chen (2023) shows that these entrepreneurial returns are capable of generating the most salient features of the wealth distribution, including the speed of wealth accumulation that leads to top-wealth holders and the distribution of capital income and returns. However, modeling entrepreneurial returns is expensive and not always fundamental for many applications in which it is needed to capture top-end wealth inequality but not necessarily the source of returns. Therefore, we consider two versions of this framework, both use a lifecycle demographic structure and a simple earnings process that abstracts from the non-Gaussian features of the second framework. The two versions differ however in their modeling of returns. The first version assumes a Markovian structure for returns, exogenous to the model. The second version is modeled after Guvenen et al. (2023) and explicitly introduces entrepreneurial activity.[3]

Our analysis provides clear and distinct results for each framework. First, *awesome-state shocks* are capable to reproduce some feature of the wealth distribution (like its Gini coefficient and its top 1% share) but are unable to generate a thick right tail. Because of this, this framework does not generate the number of "hundred-millionaires" observe in the data, misplacing the top percentiles of the wealth distribution well below their actual data levels. For instance, while the 99^{th} percentile of the US wealth distribution

kurtosis (as high as 35 compared with 3 for a Gaussian distribution) implies that in a given year most individuals experience very small changes in earnings relative to overall standard deviation, but very few experience extremely large shocks. They also find significant variation in the statistical properties of earnings shocks over the life cycle and with the income level of individuals.

[3]Boar and Midrigan (2023) extend Boar and Midrigan (2022) to include entrepreneurs within an infinite horizon setup where workers receive awesome-state shocks. Brüggemann (2021) and Brüggemann and Mahone (2025) introduce entrepreneurship into a perpetual-youth demographic structure with and without awesome-state shocks, respectively.

is $3.5 million (Smith, Zidar, and Zwick, 2022), the framework places it at $1.7 million. Thus, the framework captures relative moments (like wealth shares) while missing the level of the distribution.

Moreover, the assumption of a *perpetual-youth demographic structure* generates counterfactual implications for the distribution of wealth over age, and indeed for the age distribution of agents implied by the model. Realistic probabilities of retirement and death (working on average 45 years and dying 15 years after retirement) lead to 27 percent of a birth-cohort being alive at age 100 with over 18 percent of them still working (by age 200 there 2.3 percent of the cohort is still alive). This demographic structure has strong implications for the distribution of wealth: wealthy agents are old-working agents. In fact, centenarians hold 30–40 percent of all the wealth in the economy. As before, this framework matches wealth concentration but misplaces wealth holders by over-representing (extremely) old individuals in the top of the distribution.

Second, a *empirically grounded earnings* and aging process is not capable (on its own) to reproduce any of the features of top wealth inequality. In particular it does not generate high wealth concentration at the top, producing a top 1% wealth share of 13 percent (compared to 31 percent in the data) even though it produces the same 99^{th} percentile as models of the first framework ($1.7 million). This framework is also unable to generate a thick right tail of the wealth distribution with no individual holding more than $20 million, consistent with theoretical results highlighted in the literature (e.g., Benhabib and Bisin, 2018; Stachurski and Toda, 2019).

Third, *heterogeneity in rates of return* reproduces the main features of the wealth distribution, including the shape of the right tail, the prevalence of hundred-millionaires, and top wealth concentration. The version of this framework that explicitly models *entrepreneurial returns* performs best, being able to also capture the speed of wealth accumulation that leads to self-made billionaires who make up over half of the richest individuals in the economy. Notably, the much simpler version with *Markovian returns* performs just as well in terms of matching wealth concentration statistics, the placement of top percentiles, and the shape of the right tail of the distribution. The simplicity of this version comes at the cost of being unable to match the prevalence of self-made billionaires because of the lack of persistence in returns that is afforded to the entrepreneurial returns version by the introduction of permanent productivity heterogeneity between individuals.

Finally, we also use the frameworks to discuss the welfare costs of idiosyncratic shocks

by computing the consumption-equivalent welfare gain of eliminating income risk. Our results highlight that, even though the modeling of the earning process does not affect the ability to capture (or not) top wealth inequality, correctly capturing earnings dynamics remains quite relevant, in way that interacts with the mechanisms that lead to high wealth concentration.

We draw three main conclusions from this exercise: *(i)* the welfare costs of idiosyncratic fluctuations under the empirically grounded non-linear and non-Gaussian earnings process are about *two to three times* higher than under standard Gaussian processes, reaching as high as 25% to 40% of lifetime consumption; *(ii)* caution must be exercised in measuring the degree of partial insurance because different approaches yield conflicting results under non-normality; and *(iii)* consumption growth displays strong higher-order moments—negative skewness and excess kurtosis—as well as higher volatility. These results show that accounting for higher-order moments of earnings dynamics can have a first-order effect on some important economic questions in ways that interact with other motives for wealth accumulation, and they invite more work in order to understand its implications in other settings.[4]

Taken together, our results highlight the pitfalls of over-relying on matching relative moments of the wealth distribution (like top wealth concentration), as those can be matched while missing the placement and shape of the distribution as a whole (like top wealth percentiles and the thickness of the right tail). These pitfalls can lead to results in which top wealth inequality is attributed to the wrong mechanisms in ways that are inconsistent with the speed at which individuals accumulate wealth over their life cycles.

2 Frameworks for wealth accumulation

We evaluate the ability of different frameworks used in the context of incomplete-markets heterogeneous agent models to generate high levels of wealth inequality. These frameworks make use of different mechanism to achieve this goal that are a subset of the most salient mechanisms considered in the literature: *(i)* life cycle/retirements saving motives (as in the foundational work of Friedman, Ando, and Modigliani); *(ii)* buffer-stock

[4]The analysis of consumption-savings decisions in the presence of higher-order income risk has also been conducted previously in Wang (2007), and recently in Wang, Wang, and Yang (2016) and De Nardi, Fella, and Pardo (2016). Our analysis complements these papers by embedding a very general stochastic process (complete with age and income dependent shocks) with minimal modifications into a life-cycle model and analyzing a range of its properties.

precautionary savings (as in Zeldes, 1989, Deaton, 1991, Carroll, 1992, and Aiyagari, 1994, among many others); *(iii)* labor income risk augmented by awesome state shocks that generates high savings rates among high earners (as in Castañeda, Díaz Giménez, and Ríos Rull, 2003); *(iv)* perpetual-youth demographic structure that leads to longer wealth accumulation periods (as in Castañeda, Díaz Giménez, and Ríos Rull, 2003); and *(v)* rate of return heterogeneity (as in the seminal work of Champernowne).[5]

The three frameworks we study build on a common model of consumption and saving choices in the presence future retirement and income risk (mechanisms *(i)* and *(ii)*), but differ in whether other mechanisms are included. The first framework combines mechanisms *(iii)* and *(iv)*, so that income risk includes awesome states of short expected duration, and the life cycle of individuals evolves stochastically, first aging into retirement and then dying. The second framework maintains the mechanisms of the common consumption-saving model, but matches the empirical features of aging and retirement in the US population (following the Social Security Administration's demographic tables) and includes an empirically motivated earnings process that matches the actual dynamics of income in the US, as described in Guvenen, Karahan, Ozkan, and Song (2021a). The third framework introduces mechanism *(v)* but simplifies the earnings process.

Model of savings Consider a continuum of dynasties that exist indefinitely. In each generation individuals participate in the labor market for the first T_W years of their life, retire at that age, and die at age $T > T_W$ (or stochastically). Let δ_h be the conditional survival probability from period h to $h+1$. When an individual dies, a newborn in the next generation replaces them and inherits their wealth and their ex-ante type (α^i, β^i), albeit imperfectly. The older generations show imperfect dynastic altruism towards their offspring, according to a warm-glove bequest function (e.g., $\phi(b) = \phi_1 (b + \phi_2)^{1-\gamma}$). All individuals have CRRA preferences over consumption alone and supply labor inelastically. Prices are taken as given by individuals and do not change over time. We simulate this model for several generations until we reach a stationary distribution of wealth.

In all generations, individual earnings follow a Markovian stochastic process whose dynamics are captured by a state variable $(z^i_{1,t})$. We solve the model under different parameterizations of the earnings process and discuss them in Section 3.2. Individuals can borrow or save using a an asset with gross return R^i_t, where borrowing is subject to an age-dependent worker-type-specific limit, denoted by $\underline{A}^k_t$. Our baseline assumes that

[5] We will also consider extensions to stochastic variations in the discount factor that lead to difference in savings rates (as in Krusell and Smith, 1997).

savings are risk-free and that $R_t^i = R$, but we also consider two alternative versions of the model where R_t^i fluctuates following an idiosyncratic Markov process and or where it corresponds to entrepreneurial returns that depend on individual entrepreneurial productivity (z_t^i) and asset holdings. We describe the setup for entrepreneurial returns in Section 3.3.

Then, the budget constraint is given by

$$\begin{aligned} c_t^i + a_{t+1}^i &= a_t^i R_t^i + Y_t^{\text{disp},i}, \qquad \forall t, && (1)\\ Y_t^{\text{disp},i} &= \lambda \max\left\{\underline{Y}, \widetilde{Y}_t^i\right\}^{1-\tau}, \qquad t = 1, \ldots, T_W, && (2)\\ Y_t^{\text{disp},i} &= \left(\widetilde{Y}_R^k\right)^{1-\tau}, \qquad t = T_W + 1, \ldots, T, && (3)\\ a_{t+1}^i &\geq \underline{A}_t^k, \qquad \forall t, && (4) \end{aligned}$$

where c_t^i and a_t^i denote consumption and asset holdings, respectively, and $Y_t^{\text{disp},i}$ is disposable income, which differs from gross income, $\widetilde{Y}_t^i$, in two ways. First, the government provides social insurance by guaranteeing a minimum level of income, $\underline{Y}$, to all individuals—more on this in a moment. Second, the government imposes a progressive tax on after-transfer income. Following Benabou (2000) and Heathcote, Storesletten, and Violante (2014), we take this tax function to have a power form, with exponent $1 - \tau$. This specification has been found to fit the U.S. data quite well. Finally, retirees receive pension income, $\widetilde{Y}_R^k$, specified to mimic the U.S. Social Security Administration's OASDI system.[6] Appendix D contains further details. The dynamic programming problem of an individual is

$$\begin{aligned} V_t^i\left(a_t^i, z_{1,t}^i, z_{2,t}^i; \Upsilon^k\right) &= (1-\beta)\frac{\left(c_t^i\right)^{1-\gamma}}{1-\gamma} + \beta(1-\delta_h)\,\mathbb{E}_t\left[V_{t+1}^i\left(a_{t+1}^i, z_{1,t+1}^i, z_{2,t+1}^i; \Upsilon^k\right)\right] \\ &\quad + \beta\delta_h \frac{\phi_1\left(b+\phi_2\right)^{1-\gamma}}{1-\gamma} \qquad (5) \\ &\text{s.t. equations (1)-(4)}, \end{aligned}$$

where β is the subjective discount factor, and γ governs risk aversion. This problem can be solved using standard numerical techniques; see computational Appendix D.

[6] The dependence of pension on worker type k rather than individual income is to avoid introducing another state variable.

3 Calibration

The models in the three frameworks we study share most of its structure. We therefore start by describing their common features and then describe in details the difference in demographic structure, earnings, and returns.

The coefficient of relative risk aversion γ is set to 2 as a conservative choice. The discount factor, β, is calibrated to replicate the wealth-to-income ratio of 4 as reported in the Flow of Funds Z1 tables.[7] We assume that there is no borrowing ($\underline{A}_t^k = 0$), consistent with standard practice in the literature (e.g., Castañeda, Díaz Giménez, and Ríos Rull, 2003; Guvenen et al., 2023). We also abstract from bequest motives in our main exercise, setting $\phi_1 = 0$. We set the interest rate R at 3 percent and the wage so that the average labor income by working households is equal to \$72.588 USD, the compensation of employees in the US in 2013 divided by the number of households.

As for taxes, the parameter determining progressivity, τ, is set to 0.185 following Heathcote, Storesletten, and Violante (2014). λ is calibrated such that average after tax-after transfer income is 80% of average before tax before transfer income. As for $\underline{Y}$, Hubbard, Skinner, and Zeldes (1995) estimate the total value of various government programs (food stamps, AFDC, housing subsidies, etc.) for a female-headed family with two children and no outside earnings and assets. They obtain a value of about \$7,000 in 1984 dollars—or about \$12,800 in 2010 dollars. Using the OECD equivalence scale for one adult plus two children, this comes to \$6,400 per adult person. However, as we shall see, the effects of higher-order income risk are mitigated as $\underline{Y}$ is increased, so for our baseline calibration we take a higher—and thus conservative—value of \$10,000 per year. (To put this number in context, the 10th percentile of the RE distribution in Section 3 is about \$9,000 per year.) For comparison, we shall also report and discuss results when $\underline{Y}$ is reduced to \$2,000. The entrepreneurial return model differs in the form of taxation, instead having linear labor income taxes of 22.4 percent, corresponding to $\lambda = 0.224$ and $\tau = 0$. It also has consumption taxes of 7.5 percent and capital income taxes of 25 percent. Estate taxes are of 40 percent in all models.

[7] The total wealth-to-income ratio is defined to be total asset holdings in the population relative to the sum of total before-tax labor income and capital income.

3.1 Demographic structure

We assume individuals in all frameworks enter the labor market at age 25. After that, their aging differs depending on the demographic structure of the framework. We consider two options:

Life-cycle dynamics Individuals age by one year in every model period. They work they work until age 65 when they retire, and die for certain at age 100. All individuals are however subject to mortality risk and face an age-dependent probability of dying every period (which gives rise to accidental bequests). Mortality risk is captured by the survival probability δ_h, obtained from the Social Security Administration Bell and Miller (2002). See Figure 7a discussed in the next Section .

Perpetual-youth dynamics Individuals transition between working-age, retirement-age, and death stochastically and with constant probabilities that are independent of their biological age. In particular, we assume that a working-age individual retires with a constant rate of $1/45$. Once retired, an individual dies with probability $1/15$. Only retirees are subject to mortality risk. These probabilities are consistent with those used in Castañeda, Díaz Giménez, and Ríos Rull (2003), Kaymak and Poschke (2016), Kaymak, Leung, and Poschke (2022), among others, as well as with the average length of a working career and posterior retirement.

3.2 Earnings processes

All frameworks incorporate fluctuations in income that lead to the presence of a precautionary savings motive for agents. We consider four different specifications for the earnings process faced by individuals. We provide an overview of these processes and then discuss their implications for income dynamics.

Awesome state process This corresponds to a combination of a standard Markovian earnings process and a model of "super-star" income, the awesome state, as introduced by Castañeda, Díaz Giménez, and Ríos Rull (2003). In particular, there are 4 states of idiosyncratic income reported in Table I together with their transition probabilities and population shares. The first 3 states correspond to the "normal" values of income and exhibit high persistence. The fourth state corresponds to super-star income and is over 350 times the level of the average normal income. This state is reached with very small probabilities and is markedly less persistent than normal income states.

TABLE I – Earnings Process from Castañeda, Díaz Giménez, and Ríos Rull (2003)

	Transition Matrix, $\Gamma_{\mathcal{EE}}(\%)$			
	To s'			
From s	$s'=1$	$s'=2$	$s'=3$	$s'=4$
$s=1$	96.24	1.14	0.39	0.006
$s=2$	3.07	94.33	0.37	0.000
$s=3$	1.50	0.43	95.82	0.020
$s=4$	10.66	0.49	6.11	80.51
	Levels			
$e(s)$	1.00	3.15	9.78	1,061.00
	Population Shares			
$\gamma^{\star}_{\mathcal{E}}(\%)$	61.11	22.35	16.50	0.0389

Notes: The table reports transition probabilities, $\Gamma_{\mathcal{EE}}$, between the four labor earnings states in Castañeda, Díaz Giménez, and Ríos Rull (2003) for working-age agents that remain working (rather than transitioning to retirement). All agents transition to retirement with a probability of 2.3%, so rows sum to 97.7%. The table also reports the level of each earnings state, $e(s)$, and their corresponding population shares in the stationary equilibrium, $\gamma^{\star}_{\mathcal{E}}$.

Awesome state+ process We extend the income process described in Table I by adding an additional normal state and an additional awesome state. This makes up a 6-state Markov chain where the he first 4 correspond to normal income and the last 2 to awesome states. The additional awesome state level gives the name to this process and it allows it to capture persistence and wealth growth among top earners. This process has been adopted widely in the literature following Castañeda, Díaz Giménez, and Ríos Rull (2003), for example in Kaymak and Poschke (2016), Kaymak, Leung, and Poschke (2022), and Kindermann and Krueger (2022). The transition between normal states follows an estimation of an AR(1) process from the PSID, setting $\rho = 0.985$.

Full earnings process We also consider a non-linear and non-gaussian income process for income that captures the main facts of income dynamics in US data (Guvenen et al., 2021a). Although this process has many parameters, all dynamics are captured through only one state variables ($z^i_{1,t}$), which makes it relatively straightforward to embed it into a dynamic programming problem. Let $\Upsilon^i \equiv (\alpha^i, \beta^i)$ denote the parameters capturing fixed (ex ante) heterogeneity. We discretize over Υ^i where each grid point corresponds to a worker type k. Therefore, some aspects of an individual's problem depends on his (discrete) ex ante type k, whereas others—including the dynamics of income—are drawn from continuous distributions that are also fully individual specific.

We provide a full description of this process in Appendix A. We simulate the model using the process and parameters taken from column (6) of Table VI, borrowed from *Model 8* of Table IV of Guvenen et al. (2014).

AR-Gaussian process Finally, we consider a simplified income process in the two versions of the return heterogeneity framework. Specifically, we adopt a standard log-normal income process,

$$\log \tilde{Y}_t^i = \log W + \underbrace{\kappa^i}_{\text{permanent}} + \underbrace{g\left(t\right)}_{\text{life cycle}} + \underbrace{e_t^i}_{\text{AR(1)}}, \tag{6}$$

for $h < T_W$, where κ^i is an individual fixed effect, $g(t)$ is a polynomial that captures the lifecycle component that is common to all individuals, and e_t^i follows an AR(1) process:

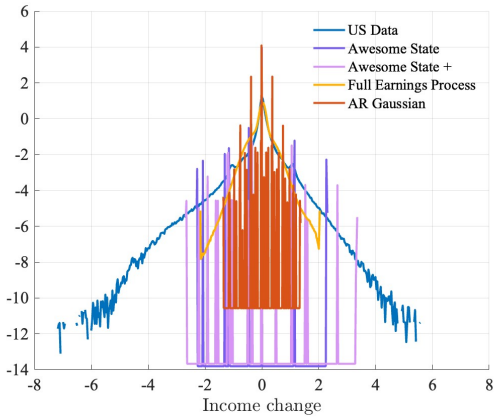
$$e_t^i = \rho_e e_{t-1}^i + \nu_t^i, \tag{7}$$

with $\nu_t^i \sim \mathcal{N}\left(0, \sigma_\nu^2\right)$. There is no inter-generational correlation between the permanent component of labor productivity, κ^i. The deterministic lifecycle profile, $g\left(t\right)$, is a quadratic polynomial that generates a 50% rise in average labor income from age 25 to its peak at age 51. The AR(1) process has a persistence of $\rho_e = 0.9$ and an innovation standard deviation of $\sigma_\nu = 0.2$, consistent with the estimates in the literature when a separate transitory shock is not modeled. We set $\kappa^i \sim N(0, \sigma_\kappa)$ to match a cross-sectional standard deviation of log labor earnings of 0.80 (Guvenen, Karahan, Ozkan, and Song, 2021b). We discretize this process for numerical exercises.

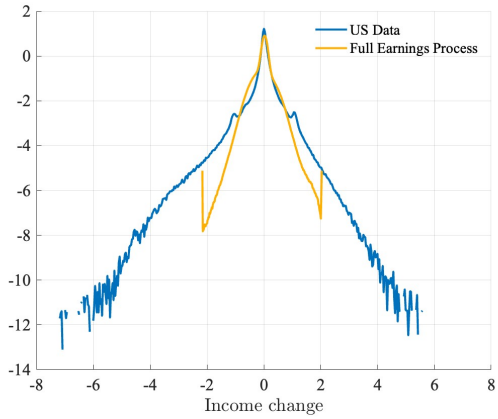
Income dynamics across earnings processes Figure 2 shows the distribution of (log) income changes in US data and contrasts it those generated with the four earnings processes just described. For scale, the figure plots the logarithm of the distribution. The distribution generated by the different earnings processes fails to fully capture that of the data, although the full earnings process (in yellow) tracks the data closely. This is evidenced in its ability to match first four moments of income changes (conditional on recent income), reported in Figure 3, like the negative skewness of income changes that Lise (2012) highlights as one of the features of the income process that leads to higher wealth accumulation. The other income processes fail at this.

The differences between earnings processes in terms of their implied higher order moments is instructive of how they affect the saving behavior of individuals. Take the

FIGURE 2 – Distribution of One Year Income Changes



(A) Log of Log-Percent Changes



(B) Log of Log-Percent Changes

Notes: Panel 3a plots the log-histogram of the log-changes in earnings one period ahead.

relative variability of income changes for low- and high-earners. While the data clearly shows that low-earners face more variability in income than high-earners, the awesome-state income process implies the opposite (and the AR Gaussian process implies no changes in variability). The awesome-state process also greatly over-estimates the kurtosis of income changes and implies counterfactually high skewness for low-earners. Taken together these moments provide a rationale for high savings among top earners that face more volatile income and negative skewness in income changes. Figure 4 reinforces this idea by plotting the evolution of income for those at the very top of the income distribution. These are not features of the changes in income among US workers.

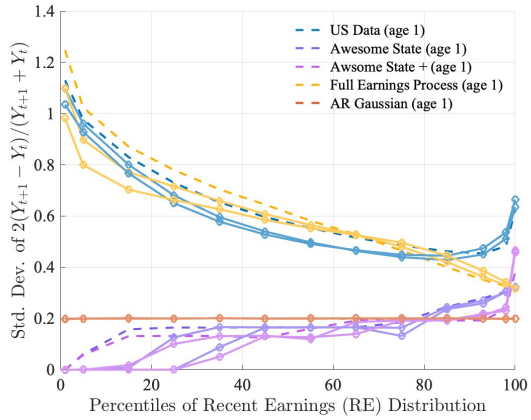
Appendix B presents additional results on the distribution of income changes.

3.3 Returns

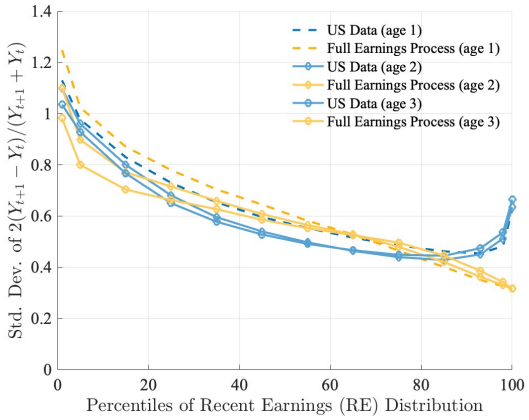
The returns on savings in the first two frameworks are assumed to be constant and equal to the risk-free rate R. Thus, individuals have homogeneous returns in the models with the awesome state, awesome state+, and full earnings processes and wealth accumulation obeys standard precautionary savings motives to guard against income fluctuations.

We consider two alternative specifications of returns for the models in the third framework with rate of return heterogeneity.

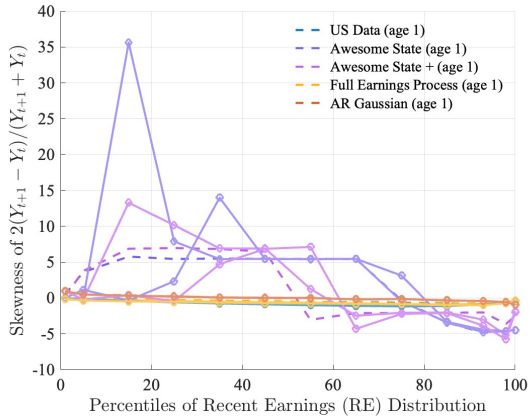
FIGURE 3 – Moments of the Distribution of 1 Year Income Changes



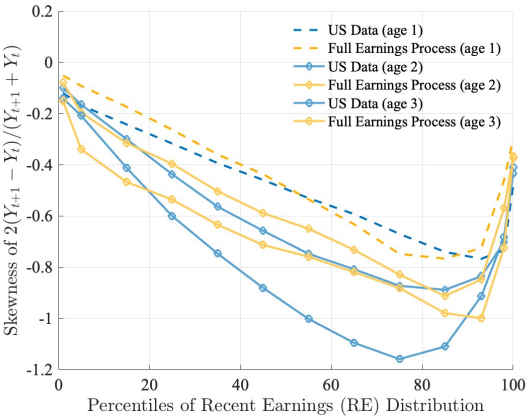
(A) Standard Deviation of One Year Changes



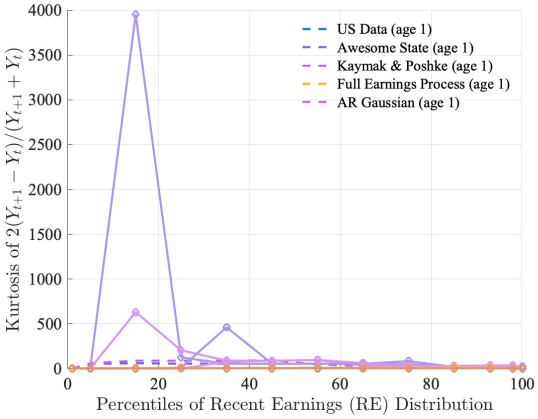
(B) Standard Deviation of One Year Changes



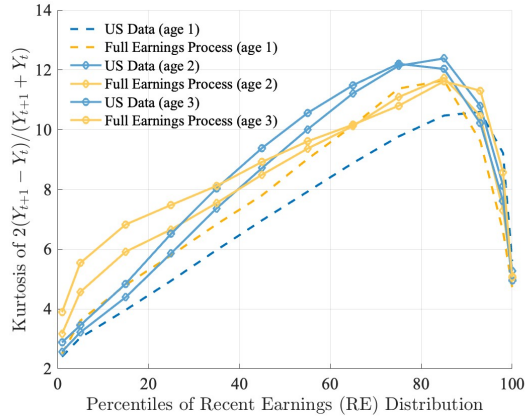
(C) Skewness of One Year Changes



(D) Skewness of One Year Changes



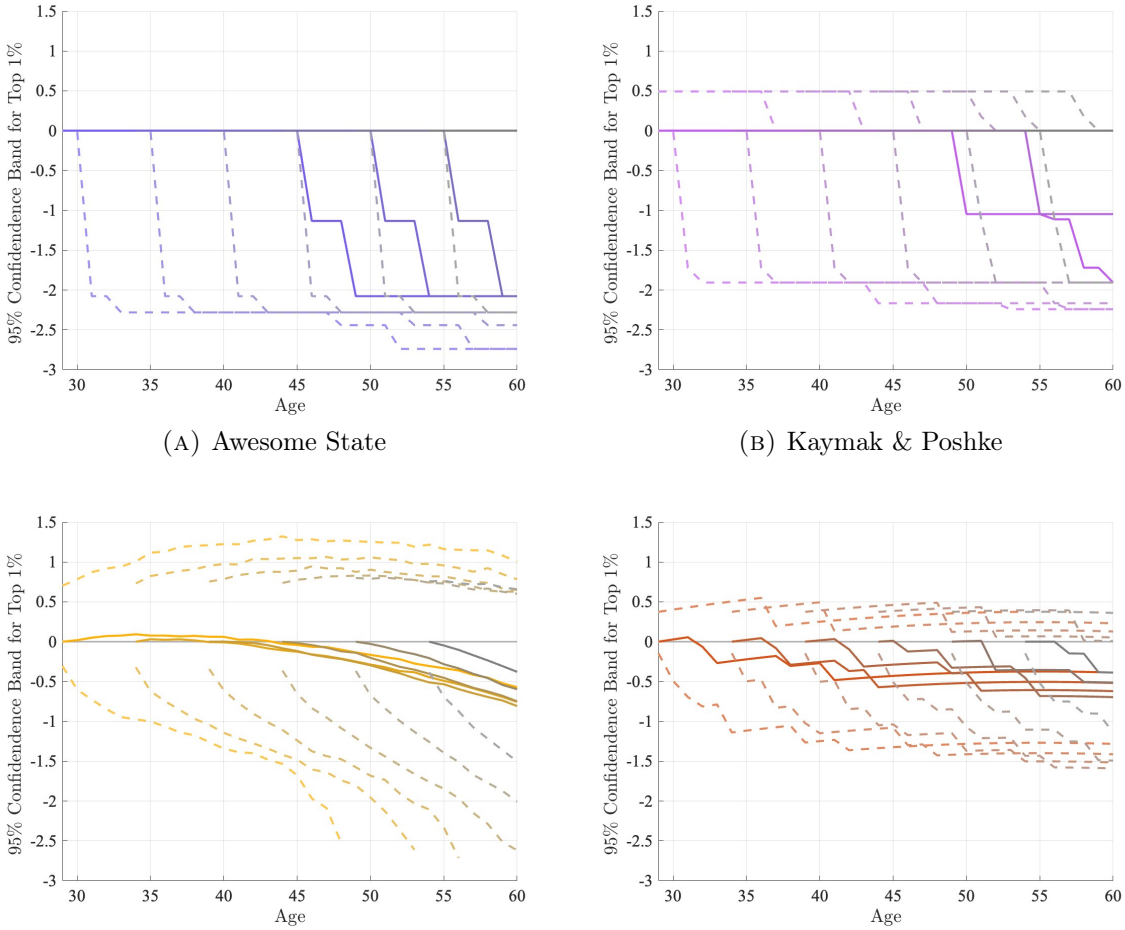
(E) Kurtosis of One Year Changes



(F) Kurtosis of One Year Changes

Notes: Panels 4a and 4b plot the standard deviation of 1 year changes in income. Panels 4c and 4d plot the skewness of 1 year changes in income. Panels 4e and 4f plot the kurtosis of 1 year changes in income. All figures include three lines for each model that correspond to different age groups. The first age group has a dashed line, the second a diamond marker, and the third a circular marker.

FIGURE 4 – Persistence at the Top of the Income Distribution



(A) Awesome State

(B) Kaymak & Poshke

(C) Full Earnings Process

(D) Discretized AR Gaussian Process

Markovian returns We allow for returns to depend on a stochastic variable z_i that follows a Markov chain with 11 states, so that $R_i = R \times \exp(z_i)$. We obtain the transition matrix and state levels of z_i by means of Rowenhurst's method (Kopecky and Suen, 2010) setting the persistence and standard deviation of the process so as to match the top 1% wealth share and the average return to wealth.

The first row of Table II reports statistics for the distribution of value-weighted returns in the model and compares it with those of the entrepreneurial returns and those of S-corporations in the US. The Markovian process for returns implies less dispersion than that observed for S-corporations, in line with evidence that these firms have high and very volatile returns. This includes the lower end of the return distribution.

TABLE II – Return Heterogeneity

	Average	p10	p50	p90
Markovian Returns	7.9	0.9	5.6	19.7
Entrepreneurial Returns	9.0	3.0	3.3	23.9
US S-Corporations	9.9	–2.8	6.2	30.1

Notes: Data for US S-Corporations comes from Smith, Zidar, and Zwick (2022). All returns are in percentage points. All statistics are wealth-weighted.

Entrepreneurial returns This version follows Guvenen et al. (2023). Returns come from the savings of individuals in risk-free bonds that pay a gross interest rate of R and of entrepreneurial profits from the operation of their businesses, π_t^i. Individuals produce differentiated intermediate goods using only capital with a linear technology, $x_t^i = z_t^i k_t^i$. Demand for each variety is iso-elastic with elasticity $\mu \leq 1$, so that the entrepreneurial production problem is

$$\pi_t^i = \max_{k_t^i \leq \vartheta\left(\overline{z}^i\right) \times a_t^i} \mathcal{P} \times \left(z_t^i k_t^i\right)^{\mu} - (R + \delta)\, k_t^i,$$

where $\mathcal{P}$ is a constant that depends on aggregate prices and R is the risk-free interest rate. For the financial constraints, we assume that they vary with the entrepreneurial productivity of individuals, so that the lowest-ability group, $\overline{z}^0$, cannot borrow at all, and the borrowing limit increases linearly with ability from there on: $\vartheta\left(\overline{z}\right) = 1 + \varphi\left(\overline{z} - \overline{z}_0\right)$. We set the value of this collateral constraint to match a debt to income ratio of 2 (consistent with the asset to income ratio of 4 targeted in all models).

The entrepreneurial *productivity* of individual i at age t, denoted z_t^i, has two components: her entrepreneurial *ability*, $\overline{z}^i$, which is a fixed characteristic of the individual, and a second component—to be described in a moment—that captures the stochastic variation in productivity over the life cycle for a given ability level. We assume that high-ability entrepreneurs (specifically, those with $\overline{z}^i > \overline{z}_{\text{median}} = 1$) start life in the *fast lane*, with positive factors amplifying their productivity above their base level, $z_t^i = \left(\overline{z}^i\right)^{\Psi}$, with $\Psi > 1$. In every subsequent year, they face the risk of losing their place in the

fast lane and dropping to their base level, $z_t^i = \overline{z}_i$, with annual probability p_1. With another probability p_2, all entrepreneurs (regardless of $\overline{z}^i$) face the risk of losing their productivity completely, $z_t^i = 0$, and "retiring" from entrepreneurship. The evolution of z_t^i can be summarized by the following three-state Markov chain, where $\mathbb{I}_t^i \in \{\mathcal{H}, \mathcal{L}, 0\}$ is an indicator function:

$$z_t^i = \begin{cases} \left(\overline{z}^i\right)^{\Psi} & \text{if } \mathbb{I}_t^i = \mathcal{H} \\ \overline{z}^i & \text{if } \mathbb{I}_t^i = \mathcal{L} \\ 0 & \text{if } \mathbb{I}_t^i = 0 \end{cases} \qquad \text{and} \qquad \Pi_{\mathbb{I}} = \begin{bmatrix} 1-p_1-p_2 & p_1 & p_2 \\ 0 & 1-p_2 & p_2 \\ 0 & 0 & 1 \end{bmatrix}, \tag{8}$$

denoting the transition matrix, with $\mathbb{I}_0^i = \mathcal{H}$ if $\overline{z}^i > \overline{z}_{\text{median}}$ and $\mathbb{I}_0^i = \mathcal{L}$ otherwise. The stochastic process for z_t^i is governed by five parameters: $\sigma_{\varepsilon_{\overline{z}}}$, Ψ, p_1, and p_2.

In order to capture the entrepreneurial choice of individuals we allow for a labor-leisure choice. We set preferences over consumption and leisure so that the period utility is

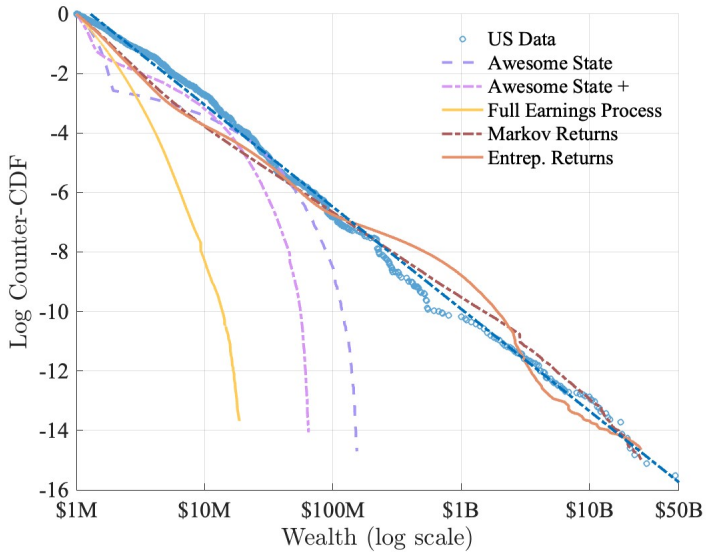
$$\frac{\left(c^{\theta}\left(1-h\right)^{1-\theta}\right)^{1-\gamma}}{1-\gamma}$$

and set θ to get average labor time to be 40 percent of total time for working-age individuals.

We set $\mathcal{P}$ so that capital income is 36 percent of total income and $\sigma_{\varepsilon_{\overline{z}}}$ to match the share of wealth of the top 1 percent of wealth holders. Finally, we set $\Psi = 5$ as in Guvenen et al. (2023) and $p_1 = 0.03$ and $p_2 = 0.01$.The model also generates a share of entrepreneurs of 10.6 percent, defined as the share of individuals with business income over half of total pre-tax income. The share of individuals whose business income is over 90 percent of total income is 4.7 percent. These individuals are 57 percent of the top 10 percent wealth holders and 79.5 percent of the top 1 percent wealth holders.

As Table II shows, the distribution of entrepreneurial returns is broadly in line with the distribution of returns for S-corporations in the US, although it does match the top and lower end of returns. This is partly because entrepreneurs in the model do not face productivity risk at the time of choosing their scale of operations, preventing negative returns.

FIGURE 5 – Pareto Tails Across Models



Notes: The figure reports the right tail of different models computed from the conditional distribution of wealth censored to start at one million US dollars. The right tail is plotted as the logarithm of the counter-CDF of wealth against the logarithm of wealth. The distribution for all models is computed on wealth over 1 million dollars so that the counter-CDF at that level is 1 by construction.

4 Top wealth inequality

What are the implications of the three frameworks we consider for top wealth inequality? Our first result verifies the well known fact that models that rely on income fluctuations fail to capture the relevant dimensions of the top wealth inequality. This holds true even when the income process exhibits excess kurtosis to match relative shares of wealth concentration (as is the case with the awesome state and awesome state+ models) or when the income process captures the observed characteristics of earnings among US workers that do imply a Pareto tail in income.[8] In this sense, the life-cycle and precautionary savings motives are insufficient to amplify income inequality into the observed levels of inequality display by wealth.

A more nuanced result points at the insufficiency of relative measures of inequality, like the Gini coefficient and top wealth shares, in gauging the success of a model in

[8]The full earnings process generates a Pareto tail for the income distribution (at least over a wide range of top income levels) unlike the Gaussian process—recall Figure 9. For example, the number of individuals with $3 million in wealth is ($e^{15/10} \approx$) about 150 times larger under the full earnings model than under the Gaussian one, and the gap increases monotonically for higher income levels. This thicker tail of the income distribution translates into a thicker tail for the wealth distribution but still fails to reproduce the observed properties of the wealth distribution in the data.

capturing wealth inequality. To see this we first consider the ability of the different models to produce a wealth distribution with a thick Pareto right tail. We plot the right tail of the wealth distribution (above one million dollars) in Figure 5. Only the heterogeneous return models can match the shape and thickness of the right tail of the wealth distribution. This is despite the Awesome state models generating more relative inequality—as measured by the Gini coefficient—and matching the top 1% share of wealth, as we report in Table III.

Matching the placement of the shape and placement of the tail of the distribution is important. It speaks to the ability of the models to capture the existence of the type of extremely wealthy individuals that populate the right tail of the wealth distribution—hundred millionaires and billionaires. As Table III shows, models that rely only on income fluctuations are unable to generate these individuals. For instance the awesome state model has only 0.003% of the population above the empirical 0.01% wealth threshold (the heterogeneous return models slightly over-state this share).

Not being able to match the presence of extremely wealthy individuals is a symptom of the underlying inability of models that rely on income fluctuations for generating high wealth growth. This is evident in the fact that these models do not generate self-made billionaires (like those in Forbes 400 list). The compounding effect of returns is able to these individuals (see the second row of Table III). Without heterogeneous returns rich individuals tend to be those who inherit large fortunes and then are lucky enough to survive a long time with them.

TABLE III – Wealth Inequality Statistics

	US Data	Income Processes			Return Processes	
		Awesome State	Awesome State +	Full Process	1st Order Markov	Entrep. Returns
Wealth Gini	0.85	0.84	0.78	0.72	0.74	0.73
Self-Made Share	54%	0.4%	0.6%	1.3%	18.8%	57.5%
	Top Percentiles of the Wealth Distribution (Millions of USD)					
p90	0.62	0.59	0.57	0.51	0.83	0.87
p99	3.52	1.70	1.67	1.70	3.29	3.18
p99.9	17.2	18.2	13.9	3.61	15.5	18.1
p99.99	77.8	56.7	30.0	6.15	99.1	94.5
	Top Wealth Shares (%)					
Top 10%	65.7	71.5	64.9	54.2	60.7	59.3
Top 1%	31.5	30.0	27.7	13.5	29.1	30.0
Top 0.1%	15.0	15.5	9.3	2.5	16.6	18.6
Top 0.01%	7.0	3.3	1.7	0.4	9.5	11.4
	Share of Population with Wealth Over X (%)					
$a \geq$\$0.62M	10	9.7	9.1	7.5	14.7	16.0
$a \geq$\$3.52M	1	0.33	0.54	0.11	0.88	0.82
$a \geq$\$17.2M	0.1	0.107	0.061	0	0.087	0.110
$a \geq$\$77.8M	0.01	0.003	0	0	0.014	0.012
	Agents over 65 years of age					
Wealth Share (%)		61.2	69.3	21.0	40.3	34.1
Pop. Share in top 10%		62.8	81.0	22.2	40.4	35.9
Pop. Share in top 1%		81.1	68.2	24.2	47.8	40.4

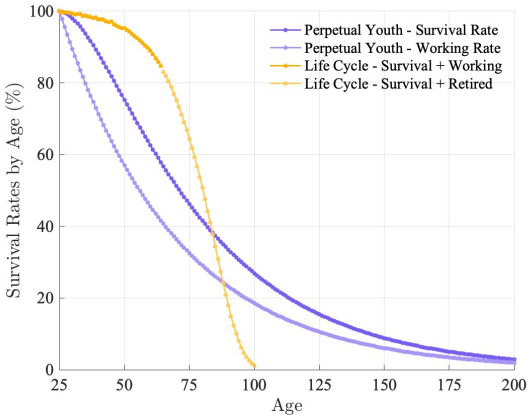
Notes: US data from Smith, Zidar, and Zwick (2022).

We next look at the demographic structure of the model and how wealth is distributed

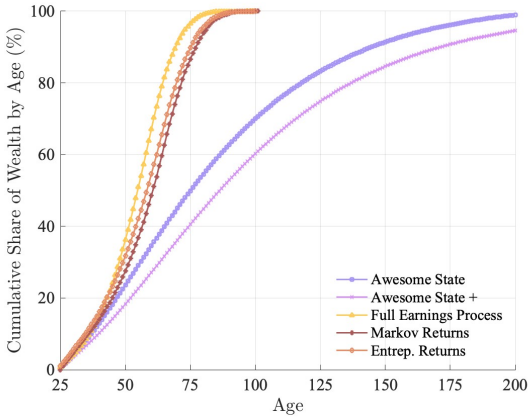
between age-groups. Figure 6 reports the cumulative survival rates by age of each model. Perpetual youth models have unrealistically high survival rates implying a population that is too old relative to what we observe in the data. The heterogeneous return model has the same survival rates as those of the US. The figure also reports the (cumulative) distribution of wealth by age, once again, the perpetual youth models with their age structure skewed toward older agents implies that too much of the wealth is held by those over 80 and 100 years of age. Finally, the demographic structure affects the life-cycle profile of wealth, implying a continuous increase in perpetual youth models.

Finally, we provide two ways of looking at how the demographic structure of the models affects the composition of top wealth groups. Table VII in Appendix C shows the distribution of age inside top wealth groups. In all models the vast majority (over 85%) of the top wealth holders are over 40 years of age, but in perpetual youth models over 20% of them are over 100 years of age. This is not the case in finite horizon models. Table VIII shows the share of wealth held by different age groups among the top wealth holders. We also report how the wealth distribution changes by age by looking at the evolution of the top percentiles and top wealth concentration by age in Figure 11, such as the age profiles of the 99^{th} of wealth and of the top 1% share of wealth.

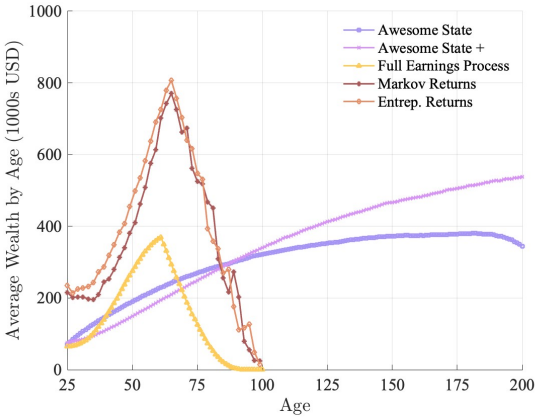
FIGURE 6 – Age Structure Across Models



(A) Stochastic Aging and Retirement vs. Life Cycle



(B) Cumulative Wealth Distribution by Age



(C) Average Wealth by Age

Notes: Panel 7a plots survival rates under stochastic aging and retirement and life cycle with survival risk. The purple lines with square markers correspond to a perpetual youth demographic structure with stochastic retirement and survival. The lighter line corresponds to the survival rate as a worker, the darker line to the survival rate. All agents start as workers and transition to retirees with a constant probability of 1/45, once retired they die with a constant probability of 1/15. Probabilities correspond to those in Castañeda, Díaz Giménez, and Ríos Rull (2003). The yellow lines with triangle markers correspond to a life cycle demographic structure with empirical survival rates from the Social Security Administration (Bell and Miller, 2002). Retirement occurs at age 65. Panel 7a plots the survival rates for a cohort born at age 20 across the different models. Panel 7b plots the CDF of wealth over age across the different models, marking the share of wealth held by agents who are of at least a given age. Panel 7c plots the average level of wealth by age across the different models.

5 Self-insurance and risk exposure

A third set of implications we explore concerns the transmission of income shocks into consumption. This transmission rate has received attention in the recent literature because it is interpreted as a measure of partial insurance—that is, insurance above and beyond self-insurance (see, e.g., Blundell, Pistaferri, and Preston, 2008, Primiceri and van Rens, 2009, and Kaplan and Violante, 2010). Importantly, this exercise highlights the importance of correctly specifying the earnings process, despite being insufficient to generate on its own the level and characteristics of top wealth inequality observed in the data.

Blundell, Pistaferri, and Preston (2008) (hereafter BPP) propose using two simple moment conditions to estimate the insurance coefficients in response to permanent shocks and transitory shocks, respectively.[9] The insurance coefficients range from 0 to 1, where 0 means no partial insurance above self-insurance (or full transmission) and 1 means perfect insurance (or no transmission). Here, the experiment we consider is the following. Suppose we give a panel data set of income and consumption simulated from the benchmark model to an econometrician and ask her to estimate the insurance coefficients for permanent and transitory shocks using BPP's moments. What would the econometrician conclude about the extent of the insurability of income shocks?

Table IV reports the results for different ages. When the true data-generating process is the NOHIP model, the BPP procedure estimates that 60% of permanent shocks are insured and the remaining 40% is transmitted to consumption at age 30.[10] The correspond-

Table IV – Measures of Insurability of Shocks

	BPP insurance coefficients					
Model:	Full Earnings Process - NOHIP			Gaussian		
Age:	30	40	50	30	40	50
Permanent	0.50	0.57	0.72	0.38	0.45	0.61
Transitory	0.95	0.94	0.94	0.97	0.96	0.96

[9] The exact formulas for the BPP insurance coefficients are $\phi^{\eta} = 1 - \frac{\text{cov}(\Delta c_t^i, y_{t+1}^{\text{disp},i} - y_{t-2}^{\text{disp},i})}{\text{cov}(\Delta y_t^{\text{disp},i}, y_{t+1}^{\text{disp},i} - y_{t-2}^{\text{disp},i})}$ for permanent shocks and $\phi^{\varepsilon} = 1 - \frac{\text{cov}(\Delta c_t^i, \Delta y_{t+1}^{\text{disp},i})}{\text{cov}(\Delta y_t^{\text{disp},i}, \Delta y_{t+1}^{\text{disp},i})}$ for transitory shocks.

[10] For brevity, we omit the statistics of the benchmark model from here on, which look very similar to the NOHIP model discussed here.

ing insurance coefficient for the Gaussian process is lower—39% (so 61% transmitted). The same gap between the insurance coefficients in the two models—of approximately 20%—is stable across different ages. Taken at face value, this finding would indicate that persistent income shocks in the benchmark model are more insurable relative to the Gaussian model. This could happen if the NOHIP *process* features less persistent shocks (which is partly true, at least judging coarsely based on the values of ρ_1 and ρ_2) or if the NOHIP *life-cycle model* features more insurance opportunities relative to the Gaussian model (which is not true—they both have self-insurance only).

However, a second way economists have approached the question of the degree of insurance is by studying how within-cohort consumption inequality evolves over the life cycle. The idea is that if shocks are easily insurable—because either they are not persistent or the economy features rich smoothing opportunities—then consumption inequality should not rise much with age. To explore this, we compute the cross-sectional variance of log consumption in the NOHIP and Gaussian models. In the NOHIP model, consumption inequality rises by about 22 log points from age 25 to age 60, which is about twice as large as the rise in the Gaussian model (of 13 log points). Therefore, this second way to look at the degree of partial insurance leads to the opposite conclusion: shocks in the NOHIP model are harder to insure, which causes consumption inequality to rise more than in the Gaussian model.[11] This latter evidence is also more consistent with the higher welfare costs of idiosyncratic shocks that we found for the NOHIP model relative to the Gaussian model above.

Taken together, we conclude from these two pieces of evidence that once we move beyond Gaussian shocks and linear models, extra care is needed to properly measure the extent of insurability of shocks. The nonlinear dynamics and higher-order moments generate interesting new patterns. For example, why does the transmission parameter indicate a low rate of transmission in the NOHIP model? An important reason is that the partial insurance coefficient measures the *average* response of consumption growth to income shocks. But it is plausible to expect that the consumption response varies by the size of the shock, by its sign, and by many of its other properties established in the previous sections. So, the average response coefficient provides an incomplete picture of

[11] Further analysis has shown that the key feature of the NOHIP model in generating the large rise in inequality is the nonlinearity and persistence of non-employment shocks. Removing the dependence of p_ν on z as we did above (and which breaks both the autocorrelation and income dependence of nonemployment) leads to a much smaller rise in consumption inequality—though still higher than the Gaussian model.

TABLE V – Cross-Sectional Moments of Consumption

Model	Full Earnings Process - NOHIP			Gaussian		
Age	30	40	50	30	40	50
Variance ($\times 10^2$)	0.92	0.70	0.30	0.80	0.70	0.30
Kelley skewness	–0.60	–0.58	-0.50	0.07	0.09	0.10
Crow-Siddiqui kurtosis	6.5	6.1	7.9	3.0	3.0	3.5

the transmission of income shocks to consumption. Instead, we would like to know the properties of the entire distribution of consumption growth rates implied by a model,[12] including its higher-order moments. We turn to this next.

Higher-Order Moments of Consumption Growth Table V reports the variance, Kelley skewness, and Crow-Siddiqui kurtosis of consumption growth in the NOHIP model and the Gaussian model. First, the variance of consumption growth declines with age in both models but is about 50% larger in the benchmark model as in the Gaussian one. Recall that both income processes have similar variances for income growth. So this suggests that consumption growth is much more volatile under the benchmark model.

Second, consumption growth is negatively skewed before age 50 in the benchmark model but has slight positive skewness in the Gaussian model. The minimum income floor induces slight positive skewness in both models—without it, the benchmark model delivers even more negatively skewed consumption growth. Interestingly, consumption growth becomes *less* negatively skewed with age, even though income shocks become *more* negatively skewed, which seems to be due to precautionary wealth allowing better smoothing at older ages.

Finally, consumption growth has very high excess kurtosis, as measured by the robust Crow-Siddiqui measure. The Gaussian model delivers almost no excess kurtosis (the Gaussian distribution has a Crow-Siddiqui kurtosis of 2.9). (The third and fourth standardized moment measures of skewness and kurtosis—which we do not report here for brevity—confirm the same finding). In further analysis, we have also found that the excess kurtosis of consumption growth increases when risk aversion is increased.

To our knowledge, there are only a few papers that have considered deviations from lognormality of the consumption growth distribution. Among these, Brav, Constan-

[12]This point is also emphasized in Arellano, Blundell, and Bonhomme (2017).

tinides, and Geczy (2002) argued that accounting for the skewness of consumption growth distribution is critical for reconciling the observed equity premium with the data, and Constantinides and Ghosh (2017) provide evidence that household consumption growth is negatively skewed in the US data. More recently, Toda and Walsh (2015) found that the household consumption growth distribution displays a double Pareto distribution, which is consistent with the implications of the benchmark model studied here displaying excess kurtosis of consumption growth. However, neither paper studies the life cycle patterns of these moments shown in Table V. An exciting future research avenue would be to confront these implications with high quality micro consumption data from which higher-order moments can be measured precisely.

6 Conclusion

In this paper we study consumption-savings implications of these new empirical findings on idiosyncratic life-cycle risk and contrast them with other frameworks that generate high levels of wealth inequality, mainly excess kurtosis coming from "awesome state" income shocks and return heterogeneity. Our analysis shows that models that rely exclusively in income fluctuations are unable to reproduce the characteristics of wealth inequality. In particular, even when earnings processes are design to capture measures of relative inequality (like the Gini coefficient), the models are unable to account for the prevalence of hundred-millionaires and billionaires and the rapid growth of fortunes of the super-wealthy within their lifetimes. Our exercises also showed that frameworks based on perpetual-youth demographic structures generate a counterfactual concentration on wealth in the hands of extremely old individuals, as they greatly overstate their population shares (by understating aging and mortality risk). By contrast, we find heterogeneous returns models to be consistent with the features of top wealth inequality, matching the levels of concentration as well as the shape and location of the wealth distribution.

References

Aiyagari, S. Rao. 1994. "Uninsured Idiosyncratic Risk and Aggregate Saving." *The Quarterly Journal of Economics* 109 (3):659–684. URL http://www.jstor.org/stable/2118417.

Arellano, Manuel, Richard Blundell, and Stéphane Bonhomme. 2017. "Earnings and Consumption Dynamics: A Nonlinear Panel Data Framework." *Econometrica* 85 (3):693–734. URL https://doi.org/10.3982/ECTA13795.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini. 2020. "Rich Pickings? Risk, Return, and Skill in Household Wealth." *American Economic Review* 110 (9):2703–47. URL https://doi.org/10.1257/aer.20170666.

Bell, Felicite C. and Michael L. Miller. 2002. "Life Tables for the United States Social Security Area: 1900-2100"." Actuarial Study 116, Office of the Actuary, Social Security Administration.

Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* 90 (1):96–129. URL https://www.aeaweb.org/articles?id=10.1257/aer.90.1.96.

Benhabib, Jess and Alberto Bisin. 2018. "Skewed Wealth Distributions: Theory and Empirics." *Journal of Economic Literature* 56 (4):1261–1291. URL https://doi.org/10.1257/jel.20161390.

Blundell, Richard, Luigi Pistaferri, and Ian Preston. 2008. "Consumption Inequality and Partial Insurance." *American Economic Review* 98 (5):1887–1921. URL https://doi.org/10.1257/aer.98.5.1887.

Boar, Corina and Virgiliu Midrigan. 2022. "Efficient redistribution." *Journal of Monetary Economics* 131:78–91. URL https://doi.org/10.1016/j.jmoneco.2022.07.003.

———. 2023. "Should We Tax Capital Income or Wealth?" *American Economic Review: Insights* URL https://doi.org/10.1257/aeri.20220192.

Brav, Alon, George M. Constantinides, and Christopher C. Geczy. 2002. "Asset Pricing with Heterogeneous Consumers and Limited Participation: Empirical Evidence." *Journal of Political Economy* 110 (4):793–824. URL https://doi.org/10.1086/340776.

Brüggemann, Bettina. 2021. "Higher Taxes at the Top: The Role of Entrepreneurs." *American Economic Journal: Macroeconomics* 13 (3):1–36. URL https://www.aeaweb.org/articles?id=10.1257/mac.20170441.

Brüggemann, Bettina and Zachary Mahone. 2025. "Entrepreneurial Rates of Return and Wealth Inequality." *Review of Economic Studies, Forthcoming* .

Cagetti, Marco and Mariacristina De Nardi. 2006. "Entrepreneurship, Frictions, and Wealth." *Journal of Political Economy* 114 (5):835–870. URL http://ideas.repec.org/a/ucp/jpolec/v114y2006i5p835-870.html.

Cagetti, Marco and Mariacristina De Nardi. 2009. "Estate Taxation, Entrepreneurship, and Wealth." *American Economic Review* 99 (1):85–111.

Carroll, Christopher D. 1992. "The Buffer-Stock Theory of Saving: Some Macroeconomic Evidence." *Brookings Papers on Economic Activity* 23 (2):61–156. URL https://www.brookings.edu/wp-content/uploads/1992/06/1992b_bpea_carroll_hall_zeldes.pdf.

Castañeda, Ana, Javier Díaz Giménez, and José Víctor Ríos Rull. 2003. "Accounting for the U.S. Earnings and Wealth Inequality." *Journal of Political Economy* 111 (4):818–857. URL https://doi.org/10.1086/375382.

Constantinides, George M. and Anisha Ghosh. 2017. "Asset Pricing with Countercyclical Household Consumption Risk." *The Journal of Finance* 72 (1):415–460. URL https://doi.org/10.1111/jofi.12471.

De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz Pardo. 2016. "The Implications of Richer Earnings Dynamics for Consumption, Wealth, and Welfare." Working paper 21917, National Bureau of Economic Research. URL https://doi.org/10.3386/w21917.

Deaton, Angus. 1991. "Saving and Liquidity Constraints." *Econometrica* 59 (5):1221. URL https://doi.org/10.2307/2938366.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2020. "Heterogeneity and Persistence in Returns to Wealth." *Econometrica* 88 (1):115–170.

Gabaix, Xavier. 2009. "Power Laws in Economics and Finance." *Annual Review of Economics* 1:255–93.

Greenwald, Daniel L, Matteo Leombroni, Hanno Lustig, and Stijn Van Nieuwerburgh. 2021. "Financial and Total Wealth Inequality with Declining Interest Rates." Working Paper 28613, National Bureau of Economic Research. URL https://doi.org/10.3386/w28613.

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen. 2023. "Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation." *The Quarterly Journal of Economics* 138 (2):835–894. URL https://doi.org/10.1093/qje/qjac047.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song. 2014. "What Do Data on Millions of U.S. Workers Say About Labor Income Risk?" Working paper, University of Minnesota.

———. 2021a. "What Do Data on Millions of U.S. Workers Reveal About Lifecycle Earnings Dynamics?" *Econometrica* 89 (5):2303–2339. URL https://doi.org/10.3982/ECTA14603.

———. 2021b. "What Do Data on Millions of U.S. Workers Say About Labor Income Risk?" *Econometrica* 89 (5):2303–2339.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2014. "Consumption and Labor Supply with Partial Insurance: An Analytical Framework." *American Economic Review* 104 (7):2075–2126. URL https://doi.org/10.1257/aer.104.7.2075.

Hubbard, R. Glenn, Jonathan Skinner, and Stephen P. Zeldes. 1995. "Precautionary Saving and Social Insurance." *Journal of Political Economy* 103 (2):360–399. URL http://www.jstor.org/stable/2138644.

Kaplan, Greg and Giovanni L. Violante. 2010. "How Much Consumption Insurance beyond Self-Insurance?" *American Economic Journal: Macroeconomics* 2 (4):53–87. URL https://doi.org/10.1257/mac.2.4.53.

Kaymak, Bariş, David Leung, and Markus Poschke. 2022. "Accounting for Wealth Concentration in the United States." Working Paper 22-28, Federal Reserve Bank of Cleveland. URL https://doi.org/10.26509/frbc-wp-202228.

Kaymak, Bariş and Markus Poschke. 2016. "The evolution of wealth inequality over half a century: The role of taxes, transfers and technology." *Journal of Monetary Economics* 77:1–25. URL https://doi.org/10.1016/j.jmoneco.2015.10.004. "Inequality, Institutions, and Redistribution" held at the Stern School of Business,New York University,April 24-25, 2015.

Kindermann, Fabian and Dirk Krueger. 2022. "High Marginal Tax Rates on the Top 1 Percent? Lessons from a Life-Cycle Model with Idiosyncratic Income Risk." *American Economic Journal: Macroeconomics* 14 (2):319–66. URL https://doi.org/10.1257/mac.20150170.

Kopecky, Karen A. and Richard M.H. Suen. 2010. "Finite state Markov-chain approximations to highly persistent processes." *Review of Economic Dynamics* 13 (3):701–714. URL https://doi.org/10.1016/j.red.2010.02.002.

Krusell, Per and Anthony A. Smith. 1997. "Income and Wealth Heterogeneity, Portfolio Choice, and Equilibrium Asset Returns." *Macroeconomic Dynamics* 1 (2):387–422. URL https://doi.org/10.1017/S1365100597003052.

Lise, Jeremy. 2012. "On-the-Job Search and Precautionary Savings." *The Review of Economic Studies* 80 (3):1086–1113. URL https://doi.org/10.1093/restud/rds042.

Ozkan, Serdar, Joachim Hubmer, Sergio Salgado, and Elin Halvorsen. 2023. "Why Are the Wealthiest So Wealthy? New Longitudinal Empirical Evidence and Implications for Theories of Wealth Inequality." Tech. rep., Working Paper. URL http://dx.doi.org/10.2139/ssrn.4362818.

Primiceri, Giorgio E. and Thijs van Rens. 2009. "Heterogeneous life-cycle profiles, income risk and consumption inequality." *Journal of Monetary Economics* 56 (1):20–39. URL https://www.sciencedirect.com/science/article/pii/S0304393208001517. Carnegie-Rochester Conference Series on Public Policy: The Causes and Consequences of Rising Economic Inequality April 25-26, 2008.

Quadrini, Vincenzo. 2000. "Entrepreneurship, Saving, and Social Mobility." *Review of Economic Dynamics* 3 (1):1–40.

Smith, Matthew, Owen Zidar, and Eric Zwick. 2022. "Top Wealth in America: New Estimates Under Heterogeneous Returns*." *The Quarterly Journal of Economics* 138 (1):515–573. URL https://doi.org/10.1093/qje/qjac033.

Stachurski, John and Alexis Akira Toda. 2019. "An impossibility theorem for wealth in heterogeneous-agent models with limited heterogeneity." *Journal of Economic Theory* 182:1–24. URL https://doi.org/10.1016/j.jet.2019.04.001.

Toda, Alexis Akira and Kieran Walsh. 2015. "The Double Power Law in Consumption and Implications for Testing Euler Equations." *Journal of Political Economy* 123 (5):1177–1200. URL https://doi.org/10.1086/682729.

Vermeulen, Philip. 2016. "How Fat is the Top tail of the Wealth Distribution?" *Reivew of Income and Wealth* .

Wang, Chong, Neng Wang, and Jinqiang Yang. 2016. "Optimal consumption and savings with stochastic income and recursive utility." *Journal of Economic Theory* 165:292–331. URL https://doi.org/10.1016/j.jet.2016.04.002.

Wang, Neng. 2007. "An equilibrium model of wealth distribution." *Journal of Monetary Economics* 54 (7):1882–1904. URL https://doi.org/10.1016/j.jmoneco.2006.11.005.

Zeldes, Stephen P. 1989. "Optimal Consumption with Stochastic Income: Deviations from Certainty Equivalence." *The Quarterly Journal of Economics* 104 (2):275–298. URL https://doi.org/10.2307/2937848.

Supplemental Appendix

A A Flexible Stochastic Process

With few exceptions, the earnings dynamics literature relies on the (often implicit) assumption that earnings shocks can be approximated reasonably well with a lognormal distribution. This assumption, combined with a linear time-series model to capture the accumulation of such shocks, allowed researchers to focus their estimation to match the covariance matrix (of log earnings), abstracting from higher-order moments. This approach ignores the empirical evidence documented in Guvenen et al. (2021a) that shows large deviations from lognormality, in which case the third and fourth moments contain valuable information.

With these considerations in mind, Guvenen et al. (2021a) follows a different approach that relies on matching the kinds of moments observed in their data set and estimate a general econometric process that has the following features: (i) an AR(1) process (z_t^i) with innovations drawn from a mixture of normals; (ii) a nonemployment shock whose incidence probability $(p_\nu^i(t, z_t))$ can vary with age or z_t or both, and whose duration (ν_t^i) is exponentially distributed; and (iii) an i.i.d. normal mixture transitory shock (ε_t^i):

$$\text{Level of earnings:} \quad \tilde{Y}_t^i = (1-\nu_t^i)e^{\left(g(t)+\alpha^i+z_t^i+\varepsilon_t^i\right)} \tag{9}$$

$$\text{Persistent component:} \quad z_t^i = \rho z_{t-1}^i + \eta_t^i, \tag{10}$$

$$\text{Innovations to AR(1):} \quad \eta_t^i \sim \begin{cases} \mathcal{N}(\mu_{\eta,1}, \sigma_{\eta,1}) & \text{with prob. } p_z \\ \mathcal{N}(\mu_{\eta,2}, \sigma_{\eta,2}) & \text{with prob. } 1-p_z \end{cases} \tag{11}$$

$$\text{Initial condition of } z_t^i\text{:} \quad z_0^i \sim \mathcal{N}(0, \sigma_{z_0}) \tag{12}$$

$$\text{Transitory shock:} \quad \varepsilon_t^i \sim \begin{cases} \mathcal{N}(\mu_{\varepsilon,1}, \sigma_{\varepsilon,1}) & \text{with prob. } p_\varepsilon \\ \mathcal{N}(\mu_{\varepsilon,2}, \sigma_{\varepsilon,2}) & \text{with prob. } 1-p_\varepsilon \end{cases} \tag{13}$$

$$\text{Nonemployment duration:} \quad \nu_t^i \sim \begin{cases} 0 & \text{with prob. } 1-p_\nu(t, z_t^i) \\ \min\{1, exp(\lambda)\} & \text{with prob. } p_\nu(t, z_t^i) \end{cases} \tag{14}$$

$$\text{Prob of Nonemp. shock:} \quad p_\nu^i(t, z_t) = \frac{e^{\xi_t^i}}{1+e^{\xi_t^i}}, \text{ where } \xi_t^i \equiv a + bt + cz_t^i + dz_t^i t. \tag{15}$$

In equation (9), $g(t)$ is a quadratic polynomial, where $t = (age - 24)/10$ is normalized age, that captures the lifecycle profile of earnings *common* to all individuals. The random number (α^i) determines ex ante heterogeneity in the level of earnings and is drawn from a

normal distribution with zero mean and a variance to be estimated. The innovations, η_t^i, to the AR(1) component are drawn from a mixture of two normals. An individual draws a shock from $\mathcal{N}(\mu_{\eta,1}, \sigma_{\eta,1})$ with probability p_z and otherwise from $\mathcal{N}(\mu_{\eta,2}, \sigma_{\eta,2})$. Without loss of generality, we normalize η to have zero mean (i.e., $\mu_{\eta,1}p_z + \mu_{\eta,2}(1-p_z) = 0$) and assume $\mu_{\eta,1} < 0$ for identification. Heterogeneity in the initial conditions of z_t is captured by $z_0^i \sim \mathcal{N}(0, \sigma_{z_0})$. Transitory shocks, ε_t^i, are also drawn from a mixture of two normals (eq. 13), with analogous identifying assumptions (zero mean and $\mu_{\varepsilon,1} < 0$).

The last component of the earnings process—and as it turns out, a critical one—is a nonemployment shock (eq. 14) that is intended to primarily capture movements in the extensive margin. So, a worker is hit with a nonemployment shock with probability p_ν (meaning $\nu_t > 0$ in equation 9) whose duration follows an exponential distribution with mean $1/\lambda$ and is truncated at 1, corresponding to full-year nonemployment (zero annual income). This shock differs from z_t and ε_t by scaling the *level* of income—not its logarithm—which allows the process to capture the sizable fraction of workers who transition into and out of full-year nonemployment every year.[13]

None of the components introduced so far depend explicitly on age or recent earnings, whereas variation along these dimensions is a key characteristic of the empirical patterns we saw. One promising way we found for introducing such variation was by making the nonemployment incidence p_ν depend on age t and z_t through the logistic function shown in equation (15). The dependence of p_ν on z_t—which we refer to it as "state dependence"—turns out to be especially important as it *induces* persistence in nonemployment from one year to the next (despite ν_t itself being independent over time). This completes the description of the stochastic process.

Although this process has many parameters, all dynamics are captured through only one state variables ($z_{1,t}^i$), which makes it relatively straightforward to embed it into a dynamic programming problem. Let $\Upsilon^i \equiv (\alpha^i)$ denote the parameters capturing fixed (ex ante) heterogeneity. We construct a one-dimensional discrete grid over Υ^i where each grid point corresponds to a worker type k. Therefore, some aspects of an individual's problem depends on his (discrete) ex ante type k, whereas others—including the dynamics of income—are drawn from continuous distributions that are also fully individual specific.

[13] It takes a –350 log point shock to z_t or ε_t for a worker with \$50,000 earnings to drop below $Y_{\min}$. Generating such large shocks with sufficiently high frequency to match worker exit rates makes it challenging to simultaneously match the high frequency of smaller shocks.

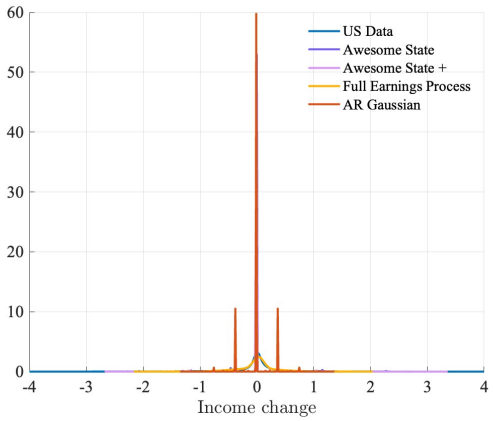
TABLE VI – Estimates of Stochastic Process Parameters

Model:	(1)	(1a)	(1b)	(3)	(7)	(8)
	Gaussian process	**Gaussian RW**	**Gaussian M7**	**Model 3**	**Model 7 NO HIP**	**Benchmark Process** Parameters
AR(1) Component	G	G	G	G	mix	mix
↪*Prob. age/income*	—	—	—	—	no/no	no/no
Nonemployment shocks	no	yes	yes	yes	yes	yes
↪*Prob. age/inc.*	—	no/no	no/no	yes/yes	yes/yes	yes/yes
Transitory shocks	G	G	G	G	mix	mix
HIP	no	no	no	no	no	yes
Parameters						
ρ	1.005	1.0	0.985	0.967	0.991	0.959
p_z					17.6%	40.7%
$\mu_{\eta,1}$					–0.524	–0.085
$\sigma_{\eta,1}$	0.134	0.155	0.210	0.197	0.113	0.364
$\sigma_{\eta,2}$					0.046	0.069
σ_{z_0}	0.343	—	—	0.563	0.450	0.714
λ				0.030	0.016	0.0001
p_ε					4.4%	13.0%
$\mu_{\varepsilon,1}$					0.134	0.271
$\sigma_{\varepsilon,1}$	0.696	0.586	0.450	0.163	0.762	0.285
$\sigma_{\varepsilon,2}$					0.055	0.037
σ_α	1.182	0.421	0.450	0.655	0.472	0.300
σ_β						0.196
$\text{corr}_{\alpha\beta}$						0.768
Objective value	74.87	73.39		35.69	27.22	22.64

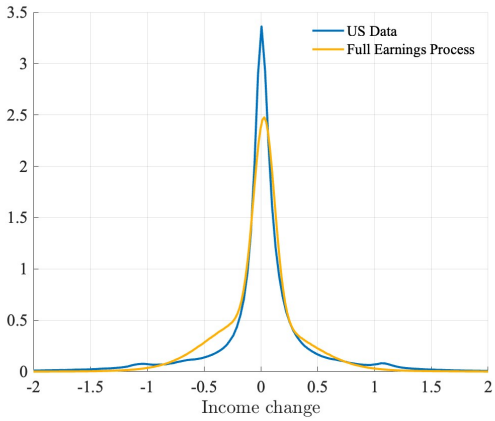
Notes: See Table 4 of Guvenen, Karahan, Ozkan, and Song (2021b).

B Extra Earnings Results

FIGURE 7 – Distribution of One Year Income Changes



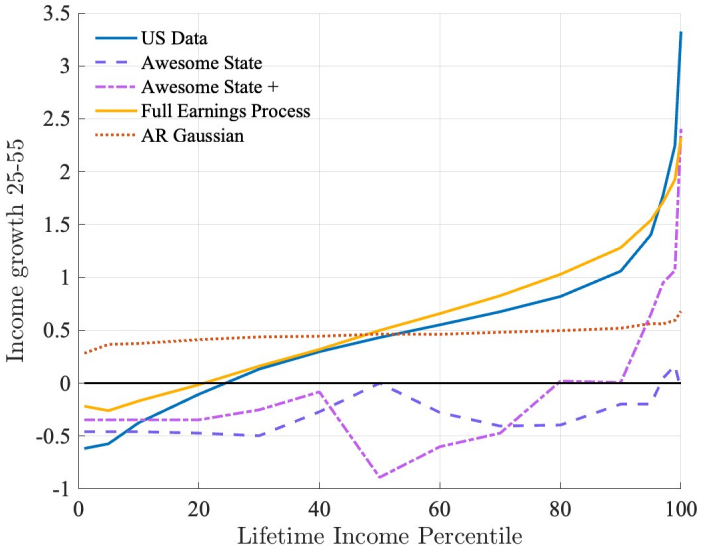
(A) Log-Percent Changes - All Models



(B) Log-Percent Changes - Full Earnings Process

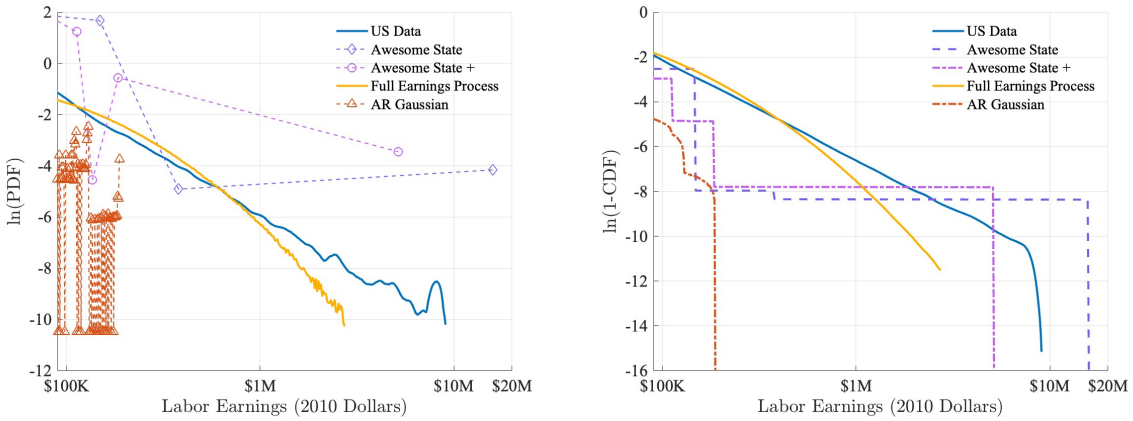
Notes: Panels 8a and 8b plot the arc-percentage changes for all models and for the full earnings process model only.

FIGURE 8 – Income Growth: Age 25–55



Notes: The figure plots the change in the natural logarithm of income between the ages of 25 and 55 for individuals in different percentiles of the distribution of lifetime income.

FIGURE 9 – Right Tail of the Income Distribution

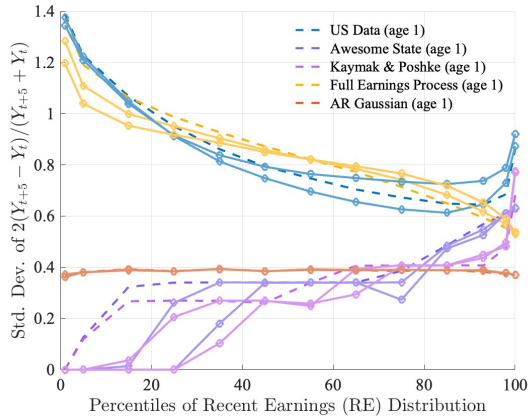


(A) Log-PDF of Income

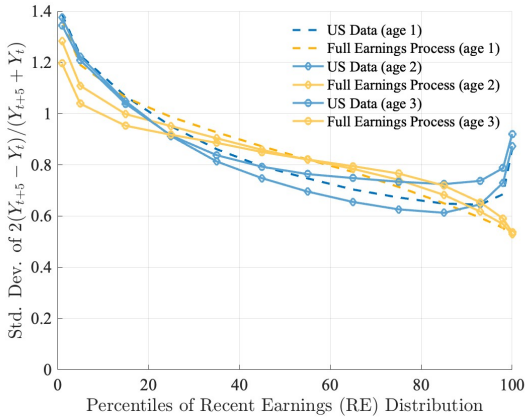
(B) Log-Counter-CDF of Income

Notes: Panel 10a plots the right tail of the log-PDF of income. The awesome state and Kaymak & Poshke lines are discrete distributions marked with diamonds and circles, respectively. Panel 10b plots the log-counter-CDF of income.

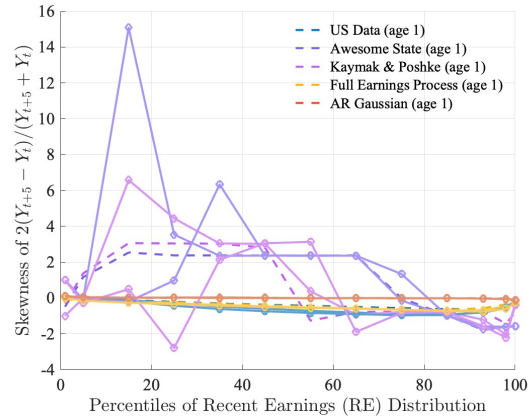
FIGURE 10 – Moments of the Distribution of 5 Year Income Changes



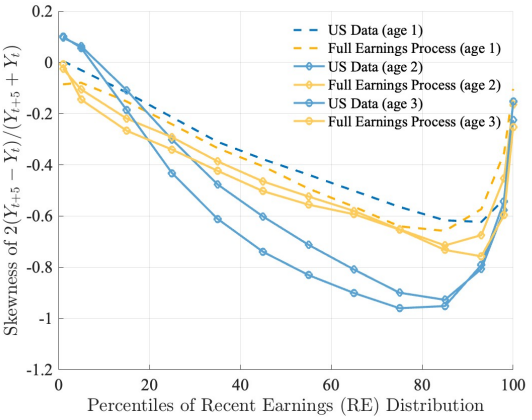
(A) Standard Deviation of Five Year Changes



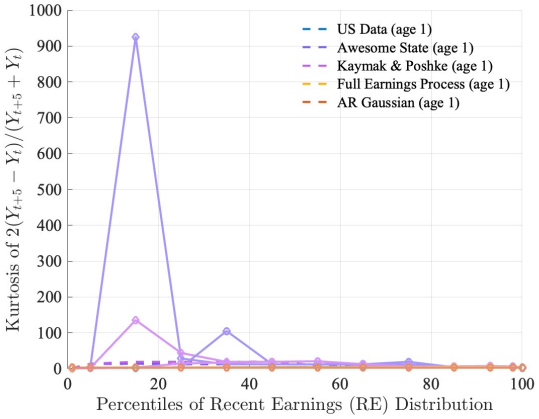
(B) Standard Deviation of Five Year Changes



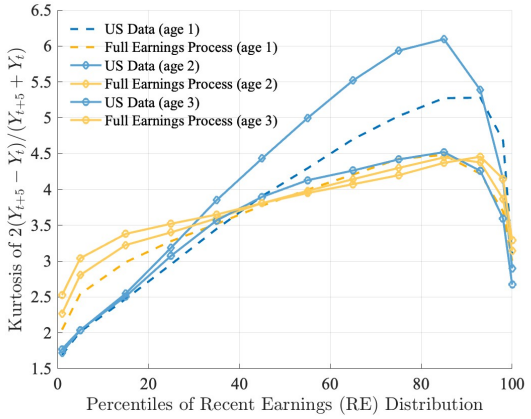
(C) Skewness of Five Year Changes



(D) Skewness of Five Year Changes



(E) Kurtosis of Five Year Changes



(F) Kurtosis of Five Year Changes

Notes: Panels 11a and 11b plot the standard deviation of 5 year changes in income. Panels 11c and 11d plot the skewness of 5 year changes in income. Panels 11e and 11f plot the kurtosis of 5 year changes in income. All figures include three lines for the data and the full earnings process model that correspond to different age groups. The first age group has a dashed line, the second a diamond marker, and the third a circular marker.

C Extra Wealth Results

TABLE VII – Age Distribution of Top Wealth Holders

(A) Awesome State Income Process

	10%	5%	1%	0.1%	0.01%
45+	86.6	97.0	90.0	84.8	93.6
65+	62.8	76.1	81.1	62.9	69.3
85+	42.9	52.5	73.6	42.8	44.5
100+	31.3	38.4	61.2	30.6	30.7
120+	19.9	24.3	39.8	19.0	17.4

(B) Awesome State + Earnings Process

	10%	5%	1%	0.1%	0.01%
45+	94.3	94.9	85.1	94.3	99.4
65+	81.0	87.5	68.2	76.7	85.4
85+	62.2	76.1	51.8	56.0	60.8
100+	49.0	63.1	41.0	42.3	42.6
120+	34.3	46.3	29.0	28.3	28.1

(C) Benchmark Earnings Process

	10%	5%	1%	0.1%	0.01%
45+	86.7	90.9	95.4	98.3	99.2
65+	22.2	23.3	24.2	23.1	20.5
85+	0.1	0.04	0.01	0.00	0.00
100+	0	0	0	0	0
120+	0	0	0	0	0

(D) Markov Returns Process

	10%	5%	1%	0.1%	0.01%
45+	87.3	88.1	85.6	78.2	75.5
65+	40.4	42.8	47.8	49.9	50.7
85+	1.5	1.8	3.2	5.6	6.7
100+	0.00	0.00	0.00	0.01	0.05
120+	0	0	0	0	0

(E) Entrepreneurial Returns Process

	10%	5%	1%	0.1%	0.01%
45+	86.2	87.5	86.4	77.3	78.9
65+	35.9	37.4	40.4	45.5	41.6
85+	0.9	1.2	1.9	4.8	4.3
100+	0.00	0.00	0.00	0.00	0.01
120+	0	0	0	0	0

Notes: Tables report the age composition of different groups of top-wealth holders. For instance, the share of individuals who are in the top 1% of wealth holders and are over 80 years of age.

TABLE VIII – Age Composition of Wealth among Top Wealth Holders

(A) Awesome State Earnings Process

	10%	5%	1%	0.1%	0.01%
45+	87.8	91.2	85.2	86.5	94.2
65+	66.1	71.5	67.0	64.0	69.8
85+	45.8	49.9	50.8	43.0	44.7
100+	33.6	36.6	39.1	30.6	30.8
120+	21.3	23.2	25.0	18.7	17.5

(B) Awesome State + Earnings Process

	10%	5%	1%	0.1%	0.01%
45+	92.7	92.1	88.3	95.1	99.5
65+	79.1	80.1	71.2	77.8	85.7
85+	61.4	65.6	53.0	56.5	60.6
100+	48.8	53.3	41.0	42.4	42.7
120+	34.5	38.5	28.2	28.3	28.2

(C) Benchmark Earnings Process

	10%	5%	1%	0.1%	0.01%
45+	89.3	92.2	95.9	98.4	99.3
65+	22.8	23.4	24.0	22.8	20.4
85+	0.05	0.03	0.01	0.00	0.00
100+	0	0	0	0	0
120+	0	0	0	0	0

(D) Markov Returns Process

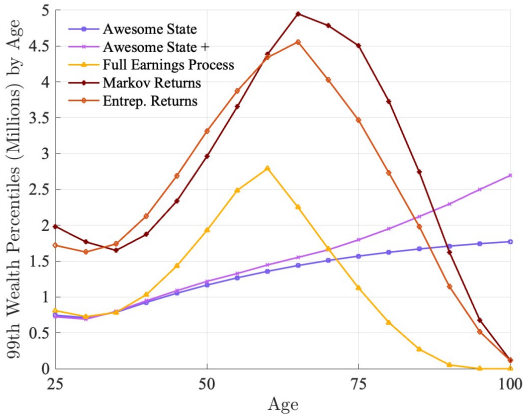
	10%	5%	1%	0.1%	0.01%
45+	84.8	84.4	81.5	78.6	78.9
65+	45.4	47.3	50.8	52.7	54.6
85+	3.2	3.7	5.2	6.7	7.4
100+	0.00	0.01	0.01	0.02	0.03
120+	0	0	0	0	0

(E) Entrepreneurial Returns Process

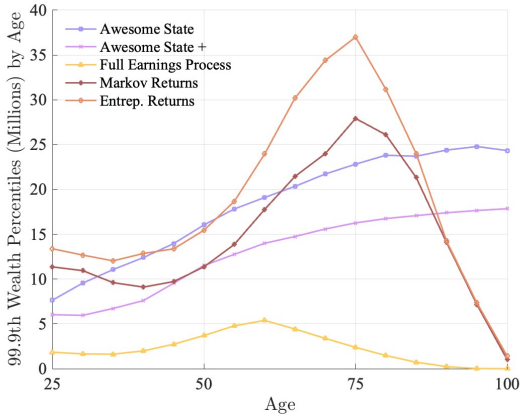
	10%	5%	1%	0.1%	0.01%
45+	83.8	83.5	80.9	78.3	77.4
65+	41.3	43.0	46.4	52.1	52.1
85+	3.0	3.5	5.0	6.9	7.7
100+	0	0	0	0	0
120+	0	0	0	0	0

Notes: Tables report the share of wealth held by given age groups among wealth holders. For instance, the share of wealth held by individuals who are in the top 1% of wealth holders and are over 80 years of age.

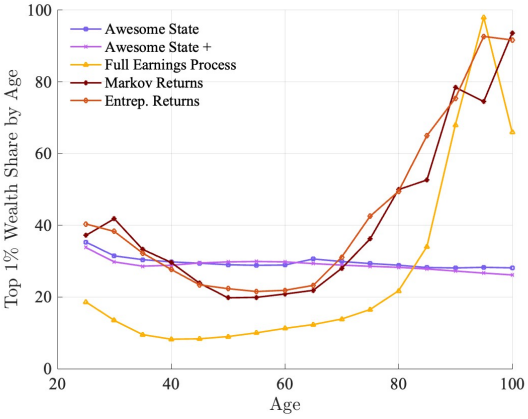
FIGURE 11 – Age-Profile of Top Wealth Across Models



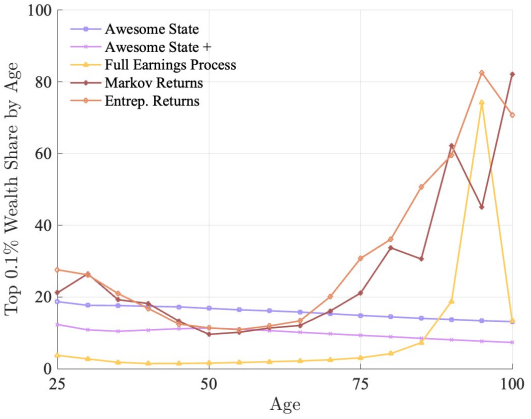
(A) 99^{th} Percentiles by Age



(B) 99.9^{th} Percentiles by Age



(C) Top 1 Percent Wealth Share by Age



(D) 0.1 Percent Wealth Share by Age

Notes: Panels 12a and 12b plot the 99^{th} and 99.9^{th} percentiles of the wealth distribution at each age bracket of 5 years. Panels 12c and 12d plot their corresponding top wealth shares. Darker colors are for the 99^{th} percentiles of wealth and the top 1% share. Lighter colors are for the 9.9^{th} percentiles of the wealth and the top 0.1% share.

D Details of Consumption Model

Retirement Pension There is a social security system that pays pension after retirement. We model the retirement pension as a function of the average earnings below social security maximum taxable earnings of worker's own type k, AE^k over the life cycle. In particular, the pension system mimics the U.S. social security "*primary insurance amount (PIA)*" which is the benefit a person would receive if he begins receiving retirement benefits at the normal retirement age:

$$y_R^k = \begin{cases} 0.9AE^k & \frac{AE^k}{AE} < 0.23 \\ 0.2AE^k + 0.32(AE^k - 0.23AE) & 0.23 < \frac{AE^k}{AE} < 1.38 \\ 0.57AE + 0.15(AE^k - 1.38AE) & \frac{AE^k}{AE} > 1.38 \end{cases}, \quad t = T_W + 1, ...T,$$

where AE is the average earnings in the economy.

D.1 Numerical Solution of the Consumption Model

In order to be able to fully capture the features of the rich earnings dynamics we estimated in this paper, we chose not to discretize the income process when solving the consumption model. Thus it is worth discussing our numerical solution methodology to solve an otherwise standard Bewley model. We employ value function iteration to solve for the policy function for the consumption decision and then use the policy function to simulate consumption-savings decisions of individuals. We now discuss the details below.

D.1.1 Bounds and Grids

For a given $\left(\alpha^k, \beta^k, \sigma_{z_1}^k, \sigma_{z_2}^k\right)$ − type worker, the value function has three continuous variables for each age t, asset holdings, and two persistent components, $a_t, z_{1,t}, z_{2,t}$. We start simulating the continuous income process which is used in the simulation of the consumption path of individuals. We take the minimum and maximum values of simulated persistent components to be the bounds of the $z_{1,t}$ and $z_{2,t}$ spaces. We choose grip points in the $z_{1,t}$ and $z_{2,t}$ spaces to be equally spaced.[14] We have 21 and 41 grid points for $z_{1,t}$ and $z_{2,t}$ spaces, respectively.

As for the asset space, the borrowing limit is set to be the each type's own simulated average earnings, $\bar{A}^k$. The upper bound for the asset spaced is chosen such that in simulations no individual's asset holdings come close to this upper bound, A_t. We choose grids in the a_t dimension to be polynomial spaced between $a_t = \bar{A}_k$ and $a_t = A_t$, with an explicit point at $a_t = 0$. We have 60 grid points for the asset space.

D.1.2 Integration and Interpolation

The value function of a type-k individual with $\left(\alpha^k, \beta^k, \sigma_{z_1}^k, \sigma_{z_2}^k\right)$ is given by:[15]

$$V_t^{i,k}\left(a_t^i, z_{1,t}^i, z_{2,t}^i\right) = \max_{c_t^i, a_{t+1}^i} u(c_t^i) + \beta \mathbb{E}_t \left[V_{t+1}^{i,k}\left(a_{t+1}^i, z_{1,t+1}^i, z_{2,t+1}^i\right) \mid z_{1,t}^i, z_{2,t}^i \right].$$

[14] Since $z_{1,t}$ and $z_{2,t}$ are in log values polynomially spaced grid points along these dimensions performed worse.

[15] The budget constraint and the other equations that properly define the value function are omitted here for simplicity.

We define the conditional expectation function:

$$\tilde{V}_{t+1}^{i,k}\left(a_{t+1}^{i},z_{1,t}^{i},z_{2,t}^{i}\right)=\mathbb{E}_{t}\left[V_{t+1}^{i,k}\left(a_{t+1}^{i},z_{1,t+1}^{i},z_{2,t+1}^{i}\right)\mid z_{1,t}^{i},z_{2,t}^{i}\right].$$

Then the value function becomes:

$$V_{t}^{i,k}\left(a_{t}^{i},z_{1,t}^{i},z_{2,t}^{i}\right)\quad=\max_{c_{t}^{i},a_{t+1}^{i}}\quad u(c_{t})+\beta\tilde{V}_{t+1}^{i,k}\left(a_{t+1}^{i},z_{1,t}^{i},z_{2,t}^{i}\right),$$

which reduces the interpolation to only along the a_{t+1} dimension, for which we use one-dimensional spline interpolation.

The conditional expectation, $\tilde{V}_{t+1}^{i,k}\left(a_{t+1}^{i},z_{1,t}^{i},z_{2,t}^{i}\right)=\mathbb{E}_{t}\left[V_{t+1}^{i,k}\left(a_{t+1}^{i},z_{1,t+1}^{i},z_{2,t+1}^{i}\right)\mid z_{1,t}^{i},z_{2,t}^{i}\right]$ involves integration over the innovations to both of the persistent components, $\eta_{1,t}^{i}$ and $\eta_{2,t}^{i}$ and an non-employment shock, ν_{t}^{i}. Non-employment shocks are functions of $\left(z_{1,t+1}^{i},z_{2,t+1}^{i}\right)$. Then for given $\left(z_{1,t+1}^{i},z_{2,t+1}^{i}\right)$ we first take the expectation over ν_{t}^{i}. For this purpose, we approximate the non-employment shocks using the Gauss-Laguerre abscissas and weights which is the quadrature that is used for exponential distribution. In particular, let $\left\{x_{\nu,t}^{1},x_{\nu,t}^{2},..x_{\nu,t}^{n_{\nu}},x_{\nu,t}^{n_{\nu}+1}=0\right\}$ and $\left\{w_{\nu,t}^{1},w_{\nu,t}^{2},..w_{\nu,t}^{n_{\nu}},w_{j,t}^{n_{\nu}+1}=1-p_{\nu,t}\right\}$ be the the Gauss-Laguerre abscissas and weights for the ν_{t}^{i}, including the case of full employment, $x_{\nu,t}^{n_{\nu}+1}=0$. Note that the weights depend on $\left(z_{1,t+1}^{i},z_{2,t+1}^{i}\right)$. Then, we define the conditional expectation function over non-employment shocks as the following:

$$\begin{aligned}\mathbb{V}_{t+1}^{i,k}\left(a_{t+1}^{i},z_{1,t+1}^{i},z_{2,t+1}^{i}\right) &= \mathbb{E}_{t,\nu}\left[V_{\nu,t+1}^{i,k}\left(\nu_{t}^{i};a_{t+1}^{i},z_{1,t+1}^{i},z_{2,t+1}^{i}\right)\mid z_{1,t+1}^{i},z_{2,t+1}^{i}\right]\\ &= \sum_{j}^{n_{\nu}+1}w_{\nu,t}^{j}V_{\nu,t+1}^{i,k}\left(x_{\nu,t}^{j};a_{t+1}^{i},z_{1,t+1}^{i},z_{2,t+1}^{i}\right).\end{aligned}$$

Innovations to the persistent components, $\eta_{1,t}$ and $\eta_{2,t}$ are drawn from a mixture of two normal distributions, one of which is a degenerate distribution. The non-degenerate normal distribution is approximated using the Gauss-Hermite abscissas and weights. Let $\left\{x_{j,t}^{1},x_{j,t}^{2},..x_{j,t}^{n_{j}},x_{j,t}^{n_{j}+1}\right\}$ and $\left\{w_{j,t}^{1},w_{j,t}^{2},..w_{j,t}^{n_{j}},w_{j,t}^{n_{j}+1}=1-p_{j,t}\right\}$ be the Gauss-Hermite abscissas and weights for the $\eta_{i,t}$ including the degenerate distribution ($x_{j,t}^{n_{j}+1}$ and $w_{j,t}^{n_{j}+1}$). Then, the conditional expectation function over the shocks to the persistent components is defined as:

$$\begin{aligned}\tilde{V}_{t+1}^{i,k}\left(a_{t+1}^{i},z_{1,t}^{i},z_{2,t}^{i}\right) &= \mathbb{E}_{t}\left[V_{t+1}^{i,k}\left(a_{t+1}^{i},z_{1,t+1}^{i},z_{2,t+1}^{i}\right)\mid z_{1,t}^{i},z_{2,t}^{i}\right]\\ &= \sum_{j_{1}}^{n_{1}+1}\sum_{j_{1}}^{n_{1}+1}w_{1,t}^{j_{1}}w_{2,t}^{j_{2}}V_{t+1}^{i,k}\left(a_{t+1}^{i},\rho_{1}z_{1,t}^{i}+x_{1,t}^{j_{1}},\rho_{2}z_{2,t}^{i}+x_{2,t}^{j_{2}}\right).\end{aligned}$$

The above part requires interpolation over $\left(z_{1,t+1}^{i},z_{2,t+1}^{i}\right)$ for which we use 2-dimensional nested cubic spline. We use 7 abscissas for innovations to the each of the persistent component.

D.1.3 Simulating the Consumption Path

The policy function for the consumption decision has three continuous variables, asset holdings, and two persistent components, $a_t, z_{1,t}, z_{2,t}$. To interpolate the consumption decision we employ the 3-dimensional nested cubic spline interpolation. We do not need an additional state for the non-employment shocks since they are transitory in nature. Namely, we store the consumption decision in the value function iteration step for only the case of full-year non-employed shocks. In simulating the non-employment shocks less than a year, we just add the difference in labor earnings due to the shorter non-employment spell to the asset holdings of the individuals, which can be thought as cash-on-hand. This allows us to interpolate the consumption decision only over three dimensions, $(a_t, z_{1,t}, z_{2,t})$ rather than 4-dimensional interpolation.

D.1.4 Other Computational Details

We use MPI for the parallel programming. In particular, we use one core for each worker type$-k$ (with a specific $\left(\alpha^k, \beta^k, \sigma^k_{z_1}, \sigma^k_{z_2}\right)$). The values of $\left(\alpha^k, \beta^k, \sigma^k_{z_1}, \sigma^k_{z_2}\right)$ and the number of individuals to be simulated for each $\left(\alpha^k, \beta^k, \sigma^k_{z_1}, \sigma^k_{z_2}\right)$ $-$type are determined using the Gauss-Hermite abscissas and weights, respectively. We have four $\left(\alpha^k, \beta^k\right)$ $-$types, three $\sigma^k_{z_1}-$types, and three $\sigma^k_{z_2}-$types. In total, we have $4 \times 3 \times 3 = 36$ ex-ante types of workers and used 36 cores to compute for 36 different value functions. Once each core finishes simulating consumption path for its own type of worker, the main core gathers all the simulated data and compute statistics for the entire economy. It takes 4-5 mins to solve and simulate the model for a given discount factor. It usually takes 6-7 iterations to converge to the discount factor that matches the wealth to income ratio observed in the data.

D.1.5 Welfare Analysis

The formula for welfare cost, χ, is given by

$$\chi = 1 - \left(\frac{V}{V_{Complete}}\right)^{1/(1-\gamma)},$$

where V is the expected lifetime utility at time 0 who has not drawn her income shocks for the first period yet, $V_{Complete}$ is the expected lifetime utility in the existence of financial markets against idiosyncratic risk (complete markets), and γ is the coefficient of relative risk aversion.

D.2 Derivation of the Risk Premium

$$U(c(1-\pi)) = \mathbb{E}\left(U(c(1+\tilde{\delta})\right)$$

Taking a first-order Taylor-series approximation to the left hand side and fourth order approximation to the right hand side we get:

$$U(c) - U'(c)c\pi = \mathbb{E}\left(U(c) + U'(c)c\tilde{\delta} + \frac{1}{2}U''(c)c^2\tilde{\delta}^2 + \frac{1}{6}U'''(c)c^3\tilde{\delta}^3 + \frac{1}{24}U''''(c)c^4\tilde{\delta}^4\right).$$

Observing that the second term on the right hand side is zero when $\mathbb{E}(\tilde{\delta}) = 0$, and rearranging yields:

$$\pi = -\frac{1}{2}\frac{u''(c)c}{u'(c)} \times m_2 - \frac{1}{6}\frac{u'''(c)c^2}{u'(c)} \times m_3 + \frac{1}{24}\frac{u''''(c)c^3}{u'(c)} \times m_4, \tag{16}$$

where m_n denotes the n^{th} central moment of $\tilde{\delta}$. To convert these into statistics that are reported in the paper, we write $m_2 = \sigma_\delta^2$, $m_3 = s_\delta \times \sigma_\delta^3$, where s_δ is the skewness coefficient, and $m_4 = k_\delta \times \sigma_\delta^4$, where k_δ is kurtosis. With this notation, and assuming a CRRA utility function with curvature γ, we get:

$$\pi^* = \frac{\gamma}{2} \times \sigma_\delta^2 - \frac{(\gamma+1)\gamma}{6} \times s_\delta \times \sigma_\delta^3 + \frac{(\gamma+2)(\gamma+1)\gamma}{24} \times k_\delta \times \sigma_\delta^4,$$

This can also be written as:

$$\pi^* = \frac{\gamma}{2} \times \sigma_\delta^2 \times \left[1 + \frac{1}{3}(\gamma+1)\left(-s_\delta \times \sigma_\delta + \frac{1}{4}(\gamma+2)k_\delta \times \sigma_\delta^2\right)\right].$$