On the Mechanics of Top Wealth Inequality*

Fatih Guvenen Sergio Ocampo Serdar Ozkan

July 2025

Abstract

Top-end wealth inequality has received 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 Aiyagari-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 either *(a)* exogenous return heterogeneity, or *(b)* endogenous entrepreneurial returns. 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: 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 framework matches many features of US income dynamics (taken from Guvenen, Karahan, Ozkan, and Song, 2021), but generates only 1/3 of the top 1% wealth share and essentially nobody with more than $100 million in wealth. By contrast, the return heterogeneity framework generate 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. It can also reproduce 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, not captured by other models.

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, University of Toronto, Federal Reserve Bank of Minneapolis, and NBER; guvenen@umn.edu; Ocampo: University of Western Ontario; socampod@uwo.ca; Ozkan: Federal Reserve Bank of St Louis, and University of Toronto; serdar.ozkan@gmail.com.

We are grateful for comments from Carter Braxton, V.V. Chari, Jonathan Heathcote, Joachim Hubmer, Rory McGee, Benjamin Moll, Emily Moschini, Elena Pastorino, and Yonseok Shin.

1 Introduction

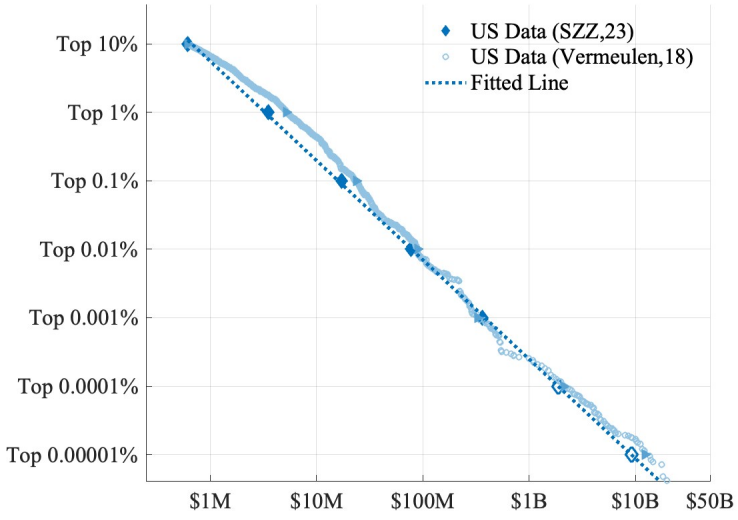
In most modern economies, wealth is extremely concentrated among a small group of households while the rest own little wealth. The right tail of the wealth distribution has a Pareto distribution with a very thick tail — implying a much higher number of super-wealthy than would be implied by lognormal 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 (e.g., Hubmer, Halvorsen, Salgado, and Ozkan, 2024). 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; Zeldes, 1989; Deaton, 1991; Carroll, 1992). 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 effects of these shocks last for a short period (5 to 10 periods on average) after which earnings fall back to their normal values. These shocks are treated as a free parameters and calibrated to match statistics

[1] For instance, Vermeulen (2018) 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: The figure plots the log of the counter CDF of the wealth distribution ($\Pr(\text{wealth} > x)$) against log wealth. Data comes from Vermeulen (2018) estimated from country level surveys merged with Forbes' billionaires list and from Smith, Zidar, and Zwick (2023) based on data from the Internal Revenue Service.

on wealth concentration. Models of this framework also use either an infinite-horizon or *perpetual-youth* demographic structure (i.e., agents age and die stochastically with constant probabilities). Examples of this framework include Kaymak and Poschke (2016), Brüggemann (2021), Greenwald, Leombroni, Lustig, and Van Nieuwerburgh (2021), 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, 2021). The process features over 20 parameters and generates *nonlinear and non-Gaussian earnings dynamics* matching 2000 empirical moments.[2] 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

[2]The non-linear and non-Gaussian properties of earnings dynamics are critical to measure the welfare costs of earnings risk and the passthrough of earnings shocks, see De Nardi, Fella, and Paz Pardo (2019), Guvenen, Ozkan, and Madera (2024), and Constantinides (2025), among others.

to insure against idiosyncratic income risk.[3]

The third framework has as its key ingredient *heterogeneity in rates of return* across households. Large and persistent return heterogeneity is 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, 2023), with entrepreneurial returns playing a central role in the wealth growth of those how make it and remain at the top of the distribution (Hubmer, Halvorsen, Salgado, and Ozkan, 2024). 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 in their modeling of returns. The first version assumes a Markovian structure for returns, exogenous to the model. The second version is taken from Guvenen et al. (2023) and explicitly introduces entrepreneurial activity.[4]

Our analysis provides clear and distinct results for each framework. First, *awesome-*

[3] Guvenen, Karahan, Ozkan, and Song (2021) 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 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.

[4] 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. Kaymak, Leung, and Poschke (2022) incorporate awesome-state shocks into a finite-life-cycle model with exogenous returns that follow a Markov process potentially correlated with the labor earnings process.

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$^{\text{th}}$ percentile of the US wealth distribution is \$3.5 million (Smith, Zidar, and Zwick, 2023), the framework places it at \$1.5 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 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 33 percent in the data) even though it produces the same 99$^{\text{th}}$ percentile as models of the first framework (\$1.5 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).[5]

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

[5]See also the quantitative exercises in De Nardi and Fella (2017) implying a limited role for earnings-risk in matching the top end of the wealth distribution. Audoly, McGee, Ocampo, and Paz Pardo (2025) show that income dynamics alone are also not enough to match the dynamics of wealth mobility along the life cycle even for individuals in the bottom 90 percent of the wealth distribution.

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. We show that introducing a small share of individuals with permanently high returns can overcome this and generate self-made billionaires while matching other aspects of the wealth distribution.

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 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).[6]

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)*,

[6]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).

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 et al. (2021). The third framework introduces return heterogeneity, mechanism *(v)*, but simplifies the earnings process. We follow the same structure as in the other frameworks for the version with exogenous returns that follow a Markov process and adopt the fully-fledged general equilibrium model of Guvenen et al. (2023) for the entrepreneurial returns version.

Model of savings[7] 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 in the perpetual-youth case). 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., $\phi(b) = \phi_1 (b + \phi_2)^{1-\gamma}$) as in De Nardi and Yang (2016). 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 focus on the stationary distribution of wealth.

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

[7] The model described here covers all of the frameworks we study except for the entrepreneurial returns model that follows instead Guvenen et al. (2023). We provide an overview in Appendix B.

worker-type-specific limit, denoted by $\underline{A}_t^k$. Then, the budget constraint is

$$c_t^i + a_{t+1}^i = a_t^i R_t^i + Y_t^{\text{disp},i}, \qquad \forall t\ ; \tag{1}$$

$$Y_t^{\text{disp},i} = \lambda \max\left\{\underline{Y}, \widetilde{Y}_t^i\right\}^{1-\tau}, \qquad t = 1, \ldots, T_W\ ; \tag{2}$$

$$Y_t^{\text{disp},i} = \left(\widetilde{Y}_R^k\right)^{1-\tau}, \qquad t = T_W + 1, \ldots, T\ ; \tag{3}$$

$$a_{t+1}^i \geq \underline{A}_t^k, \qquad \forall t\ ; \tag{4}$$

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 US data quite well. Finally, retirees receive pension income, $\widetilde{Y}_R^k$, specified to mimic the US Social Security Administration's OASDI system.[8] The dynamic programming problem of an individual is

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

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

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

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

discount factor, β, is calibrated to replicate the wealth-to-income ratio of 4 as reported in the Flow of Funds Z1 tables. 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 $\phi_1 = 0$ to focus on the role of changes in the earnings and return process. We set the interest rate R at 3 percent in all models with homogeneous returns and the wage so that the average labor income of working households is equal to \$60,462 USD, the total wages and salaries in the US in 2016 reported by the U.S. Bureau of Economic Analysis (2025) divided by the number of households reported by the U.S. Census Bureau (2025).[9]

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, 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.[10] There are also estate taxes of 40 percent in all models that dampen the intergenerational transmission of wealth. The heterogeneous return models differ in the form of taxation, instead having linear labor income taxes of 22.4 percent, corresponding to $\lambda = 0.224$ and $\tau = 0$, these models also have a less dispersed earnings process that does not generate high labor income agents, as we will see later. These models also have consumption taxes of 7.5 percent and capital income taxes of 25 percent.

[9]Differences in demographics and the average rate of return of wealth a imply differences in the average level of wealth across models. This is because of differences in the share of the population that is working (see Figure 2) that translate into differences in total income labor income and because, while all models have the same total wealth-to-income ratio, defined to be total asset holdings in the population relative to the sum of total before-tax labor income and capital income, there are differences in total capital income that reflect differences in the average rate of return to wealth.

[10]See also McGee (2021), Boar and Midrigan (2022), and Daruich and Fernández (2024) for discussion of the role of the minimum level $\underline{Y}$ for fitting tax data.

3.1 Demographic structure

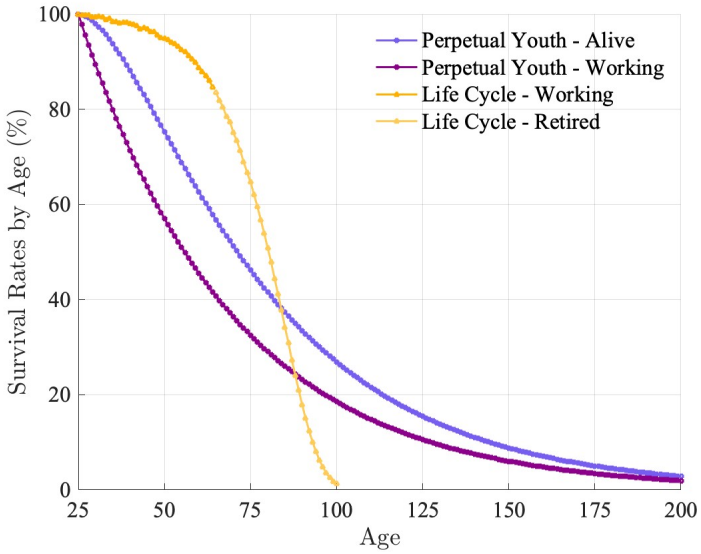
We assume individuals in all frameworks enter the labor market at age 25. After that, their aging differs depending on the demographic structure of the framework. We consider two options:

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).

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.

Implications for age-composition We next look at the demographic structure of the model and how wealth is distributed between age-groups. Figure 2 reports the cumulative survival rates by age of each model. The models with life cycle dynamics have the same survival rates as those of the US. Perpetual youth models have unrealistically high survival rates implying a population that is too old relative to what we observe in the data. The end result is an overstatement of the share of the population that is extremely old. For instance, 27 percent of the population is over 100 years old with three-quarters of them still working. Even at age 200 there is still 3 percent of the population alive, with two-thirds of them working.

FIGURE 2 – Life Cycle vs. Perpetual Youth Demographic Dynamics



Notes: Panel 2 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.

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 is the awesome state and corresponds to super-star income, over 1000 times the level of the median income. This state is reached with very small probabilities and is markedly less persistent than normal income states. This last feature turns out to be key for shaping the saving choices of individuals who find themselves in the awesome state, as they are aware that are likely to lose it.

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			
$e\left(s\right)$	1.00	3.15	9.78	1,061.00
	Population Shares			
$\gamma_{\mathcal{E}}^{\star}\left(\%\right)$	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, $e\left(s\right)$, and their corresponding population shares in the stationary equilibrium, $\gamma_{\mathcal{E}}^{\star}$.

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 $\rho = 0.985$.

Plausible Empirical Earnings Risk (PEER) 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., 2021). 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 $\Upsilon^{i} \equiv (\alpha^{i}, \beta^{i})$ 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 II, borrowed from *Model 8* of Table IV of Guvenen et al. (2014).

Models that rely only on this rich earnings process are still unable to generate the levels of top wealth concentration or the levels of top wealth observed in the data. This is despite capturing a broad set of moments mapping income dynamics conditional on age and previous income, and includes for transitory and permanent shocks as well as scarring effects of negative employment shocks. To understand better the limits of the earnings process in generating top wealth inequality we therefore consider an modified version of the PEER process that we call PEER-Top that generates a thicker tail of the income distribution with a higher variance of shocks for high-earners. This process generates top earners with over $50 million (even above what is generated by awesome state models) and it slightly overstates top incomes inequality relative to US data. Nevertheless, even this modified process is unable to replicate the main facts of top wealth inequality.

Gaussian-AR 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,

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

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

$$\eta_t^i = \rho_\eta \eta_{t-1}^i + \varepsilon_t^i, \tag{7}$$

with $\varepsilon_t^i \sim \mathcal{N}\left(0, \sigma_\varepsilon^2\right)$. There is no inter-generational correlation between the permanent component of labor productivity, α^i. We set $\alpha^i \sim N(0, \sigma_\alpha)$ to match a cross-sectional standard deviation of log labor earnings of 0.80 (Guvenen et al., 2021). 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 $\rho_\eta = 0.9$ and an innovation standard deviation of $\sigma_\eta = 0.2$, consistent with the estimates in the literature when a separate transitory shock is not modeled. We discretize this process for numerical exercises.

3.3 Return processes

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.

Markovian returns We allow for gross returns to depend on a stochastic variable z_i that follows a Markov chain with 11 states, so that $1 + R_i = \overline{R} \times \exp(z_i)$. 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.[11] We set $\overline{R} = 1.01$.

Under this framework all individuals face the same distribution of returns which, as we will see, prevents individuals born relatively poor to overcome those born wealthy, while the presence of self-made billionaires is a salient feature of the data (Guvenen et al., 2023; Hubmer et al., 2024). We therefore consider a modified version of the model that we call Markov Returns +. The objective of this second version is to show that the Markov Return process can be easily modified to match other aspects of the data (in this case the prevalence of self-made billionaires) without major changes to the rest of the framework. The modification we introduce is consequently minimal. We add a permanent type that gives individuals permanently high returns, $R_i = \hat{R}$; the returns of other individuals follow the process we just described. We calibrate this version of the model by setting the probability that a newborn born in the bottom 90 percent of the newborn wealth distribution has the new permanent type and the level of returns.[12]

[11] In the numerical implementation we start from a 15-state Markov chain and collapse the lowest 3 states resulting in 11 states that are right-skewed. This limits the extreme low-returns implied by the symmetric log-returns implied by the model and allow for high returns to take place. The resulting grid goes from -15.3 percent to 45.9 percent.

[12] This process mimics the decreasing returns to scale present in the Entrepreneurial Returns model via the exogenous change in the return process, with only initially-low-wealth individuals having access to higher returns. As in the data, this high-return type is exceedingly rare (Hubmer et al., 2024). Only 0.001 percent of newborns has this high-return type in our calibration and they have a return of $\hat{R} = 0.608$.

Entrepreneurial returns We take this framework from Guvenen et al. (2023). Unlike the previous frameworks this is a fully-fledged model in general equilibrium that incorporates other mechanisms, such as endogenous labor supply and entrepreneurship in addition to those highlighted above. We provide an overview of the key elements of the model here and a more detailed description in Appendix B. All parameter values are taken from Guvenen et al. (2023). The main aspect of this framework is that returns come from the savings of individuals in risk-free bonds that pay a gross interest rate of R (determined in equilibrium) 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

$$\pi_t^i = \max_{k_t^i \leq \vartheta\left(\overline{z}^i\right) \times a_t^i} \mathcal{P} \times \left(z_t^i k_t^i\right)^{\mu} - (R + \delta)\, k_t^i, \tag{8}$$

where $\mathcal{P}$ is a constant that depends on aggregate prices and quantities. 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 GDP ratio of 2 (consistent with the asset to income ratio of 4 targeted in all models).

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.

Audoly, Richard, Rory McGee, Sergio Ocampo, and Gonzalo Paz Pardo. 2025. "The Life-Cycle Dynamics of Wealth Mobility." NY FED Staff Report 1097, Federal Reserve Bank of New York, Staff Report. URL https://doi.org/10.59576/sr.1097.

Auerbach, Alan J, Jagadeesh Gokhale, Laurence J Kotlikoff, John Sabelhaus, and David N Weil. 1995. "The Annuitization of Americans' Resources: A Cohort Analysis." Working Paper 5089, National Bureau of Economic Research. URL http://www.nber.org/papers/w5089.

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.

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.

Braxton, J. Carter, Kyle Herkenhoff, Chengdai Huang, Michael Nattinger, Jonathan Rothbaum, and Lawrence D. W. Schmidt. 2025. "Rising Risk Among the Rich: Implications for Wealth Inequality and Interest Rates." Mimeo.

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 Dynamics, 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. 2025. "Welfare Costs of Idiosyncratic and Aggregate Consumption Shocks." *The Review of Asset Pricing Studies* 15 (2):103–120. URL https://doi.org/10.1093/rapstu/raaf002.

Daruich, Diego and Raquel Fernández. 2024. "Universal Basic Income: A Dynamic Assessment." *American Economic Review* 114 (1):38–88. URL https://doi.org/10.1257/aer.20221099.

De Nardi, Mariacristina and Giulio Fella. 2017. "Saving and Wealth Inequality." *Review of Economic Dynamics* 26:280–300. URL https://doi.org/10.1016/j.red.2017.06.002.

De Nardi, Mariacristina, Giulio Fella, and Gonzalo Paz Pardo. 2019. "Nonlinear Household Earnings Dynamics, Self-Insurance, and Welfare." *Journal of the European Economic Association* 18 (2):890–926. URL https://doi.org/10.1093/jeea/jvz010.

De Nardi, Mariacristina and Fang Yang. 2016. "Wealth inequality, family background, and estate taxation." *Journal of Monetary Economics* 77:130 – 145. URL https://doi.org/10.1016/j.jmoneco.2015.10.005.

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.

———. 2021. "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.

Guvenen, Fatih, Serdar Ozkan, and Rocio Madera. 2024. "Consumption dynamics and welfare under non-Gaussian earnings risk." *Journal of Economic Dynamics and Control* 169:104945. URL https://doi.org/10.1016/j.jedc.2024.104945. Heterogeneity and Macroeconomics of Labor Markets.

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.

Hendricks, Lutz. 2001. "Bequests and Retirement Wealth in the United States." Working paper, Arizona State University. URL https://lhendricks.org/Research/wps/bequdata_paper.pdf.

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.

Hubmer, Joachim, Elin Halvorsen, Sergio Salgado, and Serdar Ozkan. 2024. "Why Are the Wealthiest So Wealthy? New Longitudinal Empirical Evidence and Implications for Theories of Wealth Inequality." Tech. Rep. 2024-013A, Federal Reserve Bank of St. Louis, Working Paper. URL https://doi.org/10.20955/wp.2024.013.

Kaymak, Barış, 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. Forthcoming, American Economic Journal: Macroeconomcis.

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.

McGee, Rory. 2021. "Old Age Savings and House Price Shocks." Tech. Rep. 2021-4, University of Western Ontario, Department of Economics, Research Report Series. URL https://ir.lib.uwo.ca/economicsresrpt/841/.

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

Salgado, Sergio. 2020. "Technical Change and Entrepreneurship." Working paper, Wharton School at University of Pennsylvania.

Smith, Matthew, Owen Zidar, and Eric Zwick. 2023. "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.

U.S. Bureau of Economic Analysis. 2025. "Total wages and salaries, BLS [BA06RC1A027NBEA]." URL https://fred.stlouisfed.org/series/BA06RC1A027NBEA. Retrieved from FRED, Federal Reserve Bank of St. Louis, 2025-07-11.

U.S. Census Bureau. 2025. "Total Households [TTLHH]." URL https://fred.stlouisfed.org/series/TTLHH. Retrieved from FRED, Federal Reserve Bank of St. Louis, 2025-07-11.

Vermeulen, Philip. 2018. "How Fat is the Top Tail of the Wealth Distribution?" *Review of Income and Wealth* 64 (2):357–387. URL https://doi.org/10.1111/roiw.12279.

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 PEER Model: A Flexible Stochastic Process for Earnings

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. (2021) 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. (2021) 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 II – Estimates of Stochastic Process Parameters

Model:	(1) **Gaussian process**	(1a) **Gaussian RW**	(1b) **Gaussian M7**	(3) **Model 3**	(7) **Model 7 NO HIP**	(8) **PEER Model 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 (2021).

PEER-Top Model We consider an alternative version of the PEER Model that generates higher income inequality and higher concentration at the top. We call this version PEER-Top and is equivalent to the PEER model except for its treatment of high-earners (the highest ex-ante type, k). We inflate the standard deviation of persistent shock for agents by 1.5. and the standard deviation of their transitory shocks by 5. This is consistent with tax data revealing an increase in the income risk among high earners evidences by a trend of higher variability of their permanent and transitory income shocks (Braxton, Herkenhoff, Huang, Nattinger, Rothbaum, and Schmidt, 2025).

B Entrepreneurial Returns Model

We highlight the main features of the model and refer the reader to Guvenen et al. (2023) for a more detailed description. The main force driving returns is the evolution of entrepreneurial productivity. 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:

$$z_t^i = \begin{cases} \left(\overline{z}^i\right)^{\Psi} & \text{if } \mathbb{I}_t^i = \mathcal{H} \\ \overline{z}^i & \text{if } \mathbb{I}_t^i = \mathcal{L} \\ 0 & \text{if } \mathbb{I}_t^i = 0 \end{cases} \qquad \text{and} \qquad \Pi_{\mathbb{I}} = \begin{bmatrix} 1-p_1-p_2 & p_1 & p_2 \\ 0 & 1-p_2 & p_2 \\ 0 & 0 & 1 \end{bmatrix}, \tag{16}$$

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. We set $\sigma_{\varepsilon_{\overline{z}}}$ 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 $p_2 = 0.01$.

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. The model 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, consistent with Cagetti and De Nardi (2006) and Salgado (2020). 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.

There are also small differences in the preferences and demographics. Following De Nardi and Yang (2016), we choose χ and b_0 to match a bequest-to-wealth ratio of 1.18 and the 90th percentile of the bequest distribution scaled by income, which is 4.31.

De Nardi and Yang (2016) report a range from 0.88% to 1.18% for the bequest/wealth ratio in the United States (see also Hendricks, 2001 and Auerbach, Gokhale, Kotlikoff, Sabelhaus, and Weil, 1995). For comparability we adopt this same calibration for the bequest parameters in the Markov Return models. Finally, the entrepreneurial return model of Guvenen et al. (2023) uses the same survival rates as the other models (including the Markov Return models), but with life starting at age 20, giving 5 more years.